Non-linear commutativity preserving maps *

Peter Šemrl
Department of Mathematics
University of Ljubljana
Jadranska 19
SI-1000 Ljubljana
Slovenia
peter.semrl@fmf.uni-lj.si

September 15, 2005

2000 *Math. Subj. Class.* 15A27, 47B49, 51A05.

1 Introduction

We denote by M_n the algebra of all $n \times n$ complex matrices. A lot of attention has been recently paid to linear preservers, that is, linear maps on M_n that preserve a certain subset or a certain property or a certain relation (see [17, 23]). Let us mention here three classical examples: linear maps preserving rank one matrices, linear maps preserving invertibility, and linear maps preserving commutativity. The first one is important because many linear preserver problems were solved by reducing them to the problem of characterizing linear maps preservering rank one matrices. We refer to [1] for a survey on Kaplansky's problem of characterizing linear maps preserving invertibility. The importance of the third example lies in the fact that the assumption of preserving commutativity can be considered as the assumption of preserving zero Lie products. All three types of linear preservers mentioned above have been extensively studied on matrix algebras as well as on more general rings and operator algebras.

Besides linear preservers also additive and multiplicative preservers were considered in the literature. It is much more surprising that in some cases we can get nice structural results on preservers with no additional algebraic structure. Already in the forties Hua initiated the study of bijective maps (no

*Supported in part by a grant from the Ministry of Science of Slovenia

linearity was assumed) on vector spaces of matrices that strongly preserve adjacent pairs of matrices [8]-[15]. Recall that two matrices A and B are adjacent if rank $(A-B)=1$. In particular, he proved that up to a translation such maps are necessarily semilinear. For some recent improvements of this result we refer to [22, 24, 25, 26]. The problem of characterizing linear invertibility preserving maps is closely related to the problem of characterizing linear spectrum preserving maps. Here, the non-linear setting is much more interesting. Namely, there are many spectrum preserving maps that are far from being semilinear or even additive. Just choose for every $A \in M_n$ an invertible matrix T_A and define $\phi : M_n \to M_n$ by $\phi(A) = T_A A T_A^{-1}$, $A \in M_n$. Then clearly, ϕ preserves the spectrum, that is, $\sigma(\phi(A)) = \sigma(A)$, $A \in M_n$. Baribeau and Ransford proved the surprising result stating that every spectrum-preserving C^1-diffeomorphism of M_n is of this form [2].

In this paper we will study non-linear commutativity preserving maps on M_n. A map $\phi : M_n \to M_n$ preserves commutativity if $\phi(A)\phi(B) = \phi(B)\phi(A)$ whenever $AB = BA$, $A, B \in M_n$. If ϕ is bijective and both ϕ and ϕ^{-1} preserve commutativity then we say that ϕ preserves commutativity in both directions.

The main result of the paper states that if $\phi : M_n \to M_n$ is a bijective continuous map preserving commutativity in both directions, then there exist an invertible matrix T and for every $A \in M_n$ a polynomial p_A such that either $\phi(A) = Tp_A(A)T^{-1}$, $A \in M_n$, or $\phi(A) = Tp_A(A^t)T^{-1}$, $A \in M_n$, or $\phi(A) = Tp_A(\overline{A})T^{-1}$, $A \in M_n$, or $\phi(A) = Tp_A(A^*)T^{-1}$, $A \in M_n$. We also study commutativity preserving maps without the continuity assumption. Let f be an automorphism of the complex field. For every $A = [a_{ij}] \in M_n$ we denote $A_f = [f(a_{ij})]$. Then the maps $\phi, \psi : M_n \to M_n$ defined by $\phi(A) = Tp_A(A_f)T^{-1}$ and $\psi(A) = Tp_A(A_f^t)T^{-1}$, $A \in M_n$, preserve commutativity. We will give examples of bijective maps on M_n preserving commutativity in both directions that are not of one of these two simple forms. However, there is a large subset $\mathcal{C} \subset M_n$ which is invariant under every bijective map ϕ on M_n preserving commutativity in both directions and the restriction of ϕ to this subset is of one of these two nice forms.

The main tool in the proof is the characterization of bijective maps defined on rank one idempotents that preserve orthogonality in both directions. This result, related to some problems in quantum mechanics, will be extended to the infinite-dimensional case.

As an application we obtain non-linear generalizations of the structural results for Lie automorphisms of matrix algebras and the algebra $B(X)$ of all bounded linear operators on a Banach space X.

2 Statement of main results

The study of linear commutativity preserving maps on M_n started with Watkins in [29]. If $n \geq 3$, then every bijective linear commutativity preserving map ϕ on

M_n is of one of the two standard forms: $\phi(A) = cTAT^{-1} + f(A)I$, $A \in M_n$, or $\phi(A) = cTA^tT^{-1} + f(A)I$, $A \in M_n$. Here, c is a nonzero complex number, T an invertible matrix, and f any linear functional on M_n. Let us call this the classical result on commutativity preserving maps on matrix algebras. It is a special case of a much more general result on bijective linear commutativity preserving maps defined on prime algebras [3]. There exist singular linear commutativity preserving maps on M_n that are not of one of the two standard forms described above. Indeed, let $\mathcal{V} \subset M_n$ be any linear subspace of matrices such that any two members of $\mathcal{V}$ commute. Then every linear map $\phi : M_n \to M_n$ whose image is contained in $\mathcal{V}$ preserves commutativity. The long standing open problem whether every linear commutativity preserving map on M_n is either of one of the two standard forms, or maps M_n into a commutative subspace has been recently answered in the affirmative [20].

In this paper we will improve the classical result on commutativity preserving maps on matrix algebras in a differenet direction. We will consider maps on M_n that are bijective and preserve commutativity in both directions but are not assumed to be linear.

Let us start by giving some examples of such maps. Similarity transformations $A \mapsto TAT^{-1}$ and the transposition map $A \mapsto A^t$ are examples of linear bijective maps on M_n preserving commutativity in both directions. Let f be any automorphism of the complex field. Recall that the identity function and the complex conjugation are the only continuous automorphisms of the complex field but that there are also many noncontinuous automorphisms of $\mathbb{C}$ [16]. For $A = [a_{ij}] \in M_n$ we denote $A_f = [f(a_{ij})]$. The map $A \mapsto A_f$, $A \in M_n$, is a ring automorphism (bijective additive and multiplicative map) of M_n, and therefore, it preserves commutativity in both directions. But there are also many nonadditive maps $\phi : M_n \to M_n$ that preserve commutativity. To see this observe that if A and B is any pair of commuting matrices and p and q any polynomials, then $p(A)$ and $q(B)$ commute as well. Choose $p_A \in \mathcal{P}$ for every $A \in M_n$. Here, $\mathcal{P}$ denotes the set of all complex polynomials. The map $A \mapsto p_A(A)$ preserves commutativity but not necessarily in both directions. In general it is not bijective. But if it is bijective and if for every $A \in M_n$ the polynomial p_A is chosen in such a way that there exists a polynomial q_A satisfying $q_A(p_A(A)) = A$, then it preserves commutativity in both directions. This is equivalent to the requirement that A and $p_A(A)$ have the same commutant. Every map $A \mapsto p_A(A)$ which is bijective and satisfies $A' = (p_A(A))'$ will be called a regular locally polynomial map. Here, A' stands for the commutant of A.

Any composition of bijective maps preserving commutativity in both directions is again a bijective map preserving commutativity in both directions. So, at this point it would be tempting to conjecture that every bijective map $\phi : M_n \to M_n$ preserving commutativity in both directions is either of the form $\phi(A) = Tp_A(A_f)T^{-1}$, $A \in M_n$, or of the form $\phi(A) = Tp_A(A_f^t)T^{-1}$, $A \in M_n$, where T is any invertible matrix, f is any automorphism of the field $\mathbb{C}$, and $A \mapsto p_A(A)$ is a regular locally polynomial map. Although wrong this conjecture turns out to be "almost true". Namely, let define $\mathcal{C} \subset M_n$ to be the subset of all matrices $A \in M_n$ with the property that all Jordan cells in the Jordan canonical form of A are of the size 1×1 or 2×2. In other words, all zeroes of the minimal polynomial of A are either simple, or of multiplicity two. The subset $\mathcal{C}$ is rather large. In particular, it contains the set of all matrices with n distinct eigenvalues which is an open dense subset of M_n. We will prove that $\mathcal{C}$ is invariant under every bijective map ϕ on M_n preserving commutativity in both directions. Our first result states that the restriction of ϕ to this subset must be of one of the two nice forms described above.

Theorem 2.1 *Let $n \geq 3$ and let $\phi : M_n \to M_n$ be a bijective map preserving commutativity in both directions. Then there exist an invertible matrix $T \in M_n$, an automorphism f of the complex field, and a regular locally polynomial map $A \mapsto p_A(A)$ such that either $\phi(A) = Tp_A(A_f)T^{-1}$ for all $A \in \mathcal{C}$, or $\phi(A) = Tp_A(A_f^t)T^{-1}$ for all $A \in \mathcal{C}$.*

We will give an example showing that outside $\mathcal{C}$ bijective maps preserving commutativity in both directions can have a wild behaviour. However, under the additional continuity assumption we get a nice result for the whole matrix algebra.

Theorem 2.2 *Let $n \geq 3$ and let $\phi : M_n \to M_n$ be a continuous bijective map preserving commutativity in both directions. Then there exist an invertible matrix $T \in M_n$ and a regular locally polynomial map $A \mapsto p_A(A)$ such that either $\phi(A) = Tp_A(A)T^{-1}$ for all $A \in M_n$, or $\phi(A) = Tp_A(A^t)T^{-1}$ for all $A \in M_n$, or $\phi(A) = Tp_A(\overline{A})T^{-1}$ for all $A \in M_n$, or $\phi(A) = Tp_A(A^*)T^{-1}$ for all $A \in M_n$. Here, $\overline{A} = \overline{[a_{ij}]} = [\overline{a_{ij}}]$, and $A^* = \overline{A}^t$.*

The assumption that $n \geq 3$ is indispensable in the above two theorems. To see this assume that $\phi : M_2 \to M_2$ is a bijective map preserving commutativity in both directions. Then, clearly, ϕ maps the center of M_2, that is, the set of all scalar matrices, onto itself. Here, we used the term scalar matrix for any matrix λI where λ is any complex number. Observe that two nonscalar matrices $A, B \in M_2$ commute if and only if A belongs to the linear span of I and B. To verify this note that every nonscalar 2×2 matrix is either diagonalizable with two different eigenvalues, or similar to an upper triangular matrix with equal diagonal entries and the $(1, 2)$-entry equal to 1. So, a bijective map $\phi : M_2 \to M_2$ preserves commutativity in both directions if and only if it maps the set of scalar matrices onto itself and for every $A \in M_n$ we have $\phi(\text{span}\,\{I, A\}) = \text{span}\,\{I, \phi(A)\}$.

A similar result for bijective maps preserving commutativity in both directions on hermitian matrices was proved in [19]. The case of hermitian matrices is much easier since every hermitian matrix is diagonalizable and then the structure of the commutant of any subset of hermitian matrices is easy to describe.

In particular, two hermitian matrices commute if and only if they are simultaneously diagonalizable. However, in [19] non-linear commutativity preserving maps on hermitian operators were treated also on infinite-dimensional spaces.

Some starting lemmas in this paper are based on some ideas from [5], where bijective semilinear commutativity preserving maps on matrix algebras were characterized. Extending the study of preservers from semilinear to non-linear case requires some new methods. Baribeau and Ransford [2] used analytical methods to study non-linear spectrum preserving maps. Our approach will depend on a recently obtained nonsurjective version of the fundamental theorem of projective geometry. The main tool in our proof will be a structural result for orthogonality preserving injective maps on rank one idempotents. To formulate it we need some more notation. A matrix $P \in M_n$ is called an idempotent if $P^2 = P$. Denote by $I_n \subset M_n$ the subset of all idempotents of rank one. Two idempotents $P, Q \in I_n$ are said to be orthogonal if $PQ = QP = 0$. In this case we write $P \perp Q$. We say that a subset $\{P_1, \ldots, P_k\} \subset I_n$ is orthogonal if $P_i \perp P_j$ whenever $i \neq j$. For a subset $\mathcal{S} \subset I_n$ we denote by $\mathcal{S}^\perp \subset I_n$ the subset of all rank one idempotents that are orthogonal to all members of $\mathcal{S}$. A map $\xi : I_n \to I_n$ preserves orthogonality if for every pair $P, Q \in I_n$ the relation $P \perp Q$ implies $\xi(P) \perp \xi(Q)$. If ξ is bijective and $P \perp Q \iff \xi(P) \perp \xi(Q)$, $P, Q \in I_n$, then we say that ξ preserves orthogonality in both directions.

Theorem 2.3 *Assume that $n \geq 3$ and let $\xi : I_n \to I_n$ be an injective map preserving orthogonality. Then there exists a nonsingular matrix $T \in M_n$ and a nonzero endomorphism $f : \mathbb{C} \to \mathbb{C}$ such that either*

$$\xi(P) = TP_fT^{-1}, \quad P \in I_n,$$

or

$$\xi(P) = TP_f^tT^{-1}, \quad P \in I_n.$$

In this paper we consider only complex spaces. Let us just remark that our proof of this statement works for more general fields than $\mathbb{C}$.

A map $\mu : I_n \to I_n$ preserves zero products if $\mu(P)\mu(Q) = 0$ whenever $PQ = 0$, $P, Q \in I_n$. Clearly, the assumption of preserving zero products is stronger than the assumption of preserving orthogonality. So, the immediate consequence of the above theorem is the statement that every injective zero product preserving map μ on I_n is of the form $\mu(P) = TP_fT^{-1}$, $T \in I_n$ for some invertible $T \in M_n$ and some endomorphism f of the complex field. Indeed, all we have to do is to observe that the transposition map does not preserve zero products. As shown in [27], this consequence holds true even without the injectivity assumption. So, the cost we had to pay for replacing the assumption of preserving zero products by a weaker assumption of preserving orthogonality is the additional injectivity assumption. This type of results lead to improvements of the finite-dimensional case of the classical Wigner's unitary-antiunitary theorem in quantum mechanics. We refer to [18] and [27] for more detailed explanation.

For our main purpose it would be enough to prove a slightly weaker version of the above result. We decided to include this stronger version because of being interesting in its own. And we will also extend it to the infinite-dimensional case. Let X be a Banach space. We denote by $B(X)$ the algebra of all bounded linear operators on X and by $I(X) \subset B(X)$ the subset of all rank one idempotents. The dual of X will be denoted by X' and the adjoint of $A \in B(X)$ by A'. For a nonzero $x \in X$ and a nonzero $f \in X'$ we denote by $x \otimes f$ the rank one operator defined by $(x \otimes f)z = f(z)x$, $z \in X$. Note that every bounded linear rank one operator on X can be written in this form and that $x \otimes f$ is an idempotent if and only if $f(x) = 1$.

Theorem 2.4 *Let X be an infinite-dimensional Banach space and $\xi : I(X) \to I(X)$ a bijective map preserving orthogonality in both directions. Then either there exists a bounded invertible linear or conjugate-linear operator $T : X \to X$ such that*

$$\xi(P) = TPT^{-1}, \quad P \in I(X),$$

or there exists a bounded invertible linear or conjugate-linear operator $T : X' \to X$ such that

$$\xi(P) = TP'T^{-1}, \quad P \in I(X).$$

In the second case X must be reflexive.

Let us just mention that the above theorem holds true also for real Banach spaces. In the real case the formulation is even nicer since T has to be linear. And also the proof is slightly simpler because of the well-known fact that every nonzero endomorphism of the real field is the identity. Thus, in the real case every semilinear map is automatically linear. As in the finite-dimensional case we get as a direct consequence the statement that every bijective map $\xi : I(X) \to I(X)$ preserving zero products in both directions has to be of the form $\xi(P) = TPT^{-1}$, $P \in I(X)$. This theorem was the main result in [18]. It was used as a main tool for generalizing Uhlhorn's version of Wigner's theorem. Wigner's theorem tells that every quantum mechanical invariance transformation can be represented by a unitary or an antiunitary operator on a complex Hilbert space. An equivalent form in mathematical language states that every bijective transformation on the set of all one-dimensional linear subspaces of a Hilbert space preserving the angle between every pair of such subspaces (transition probability in the language of quantum mechanics) is induced by a unitary or an antiunitary operator. Uhlhorn [28] improved this result by requiring only that the map preserves the orthogonality between one-dimensional subspaces. This can be further reformulated as a result on bijective maps on the set of all hermitian rank one idempotents preserving orthogonality. So, our theorem can be considered as a non-hermitian analogue of Uhlhorn's result. Molnár's proof of the above mentioned characterization of zero product preserving maps was rather long and involved the application of Ovchinnikov's characterization of automorphisms of the poset of idempotent operators [21]. A short proof based on a direct application of projective geometry was given in [27]. This simple proof provides also a short proof of Molnár's extension of Uhlhorn's theorem to the spaces with indefinite inner product. Here we improve this result by replacing the zero product preserving assumption by a weaker orthogonality preserving assumption. The cost for this generalization is a longer more complicated proof. Let us conclude these remarks by mentioning that this kind of results can be applied in the study of automorphisms of operator semigroups (see [27]).

The space M_n is a Lie algebra with the Lie product $[A, B] = AB - BA$. It is well-known that every Lie automorphism of M_n, that is, every bijective linear map $\phi : M_n \to M_n$ satisfying $\phi([A, B]) = [\phi(A), \phi(B)]$, $A, B \in M_n$, is either of the form $\phi(A) = TAT^{-1} + c\,\mathrm{tr}(A)I$, $A \in M_n$, or of the form $\phi(A) = -TA^tT^{-1} + c\,\mathrm{tr}(A)I$, $A \in M_n$. Here, $T \in M_n$ is an invertible matrix, $c \in \mathbb{C}$, and $\mathrm{tr}(A)$ denotes the trace of A. Obviously, preserving commutativity is the same as preserving zero Lie products. This simple observation together with Theorem 2.1 will give the following improvement of the above classical result.

Theorem 2.5 *Let $n \geq 3$ and let $\phi : M_n \to M_n$ be a bijective map satisfying $\phi([A, B]) = [\phi(A), \phi(B)]$, $A, B \in M_n$. Then there exist an invertible matrix $T \in M_n$, a scalar function φ defined on M_n satisfying $\varphi(C) = 0$ for all matrices C of trace zero, and an automorphism f of the complex field such that either $\phi(A) = TA_fT^{-1} + \varphi(A)I$ for all $A \in M_n$, or $\phi(A) = -TA_f^tT^{-1} + \varphi(A)I$ for all $A \in M_n$.*

Note that we have not assumed that ϕ is linear. Nevertheless, as a result we get the semilinearity of ϕ up to a function that maps in the center of M_n. This theorem holds true also in the case $n = 2$. The statement in this low dimensional case is even a little bit simpler. The precise formulation and the proof can be found in the eighth section. We will conclude the paper by extending this result to the infinite-dimensional case.

Theorem 2.6 *Let X be an infinite-dimensional Banach space and $\phi : B(X) \to B(X)$ a bijective map satisfying $\phi([A, B]) = [\phi(A), \phi(B)]$, $A, B \in B(X)$. Then either there exist a bounded invertible linear or conjugate-linear operator $T : X \to X$ and a function $\varphi : B(X) \to \mathbb{C}$ satisfying $\varphi([A, B]) = 0$ for every pair $A, B \in B(X)$ such that*

$$\phi(A) = TAT^{-1} + \varphi(A)I$$

for all $A \in B(X)$, or there exist a bounded invertible linear or conjugate-linear operator $T : X' \to X$ and a function $\varphi : B(X) \to \mathbb{C}$ satisfying $\varphi([A, B]) = 0$ for every pair $A, B \in B(X)$ such that

$$\phi(A) = -TA'T^{-1} + \varphi(A)I$$

for all $A \in B(X)$. In the second case X must be reflexive.

In particular, every bijective map on $B(X)$ that is a homomorphism with respect to the Lie product is automatically continuous and linear or conjugate-linear up to a scalar type function that vanishes on all commutators. In the case when X is a Hilbert space, the set of all commutators in $B(X)$ was characterized by Brown and Pearcy [4].

3 Preliminary results

Let $\mathcal{S}$ be a subset of M_n. Recall that its commutant $\mathcal{S}'$ is the space of all matrices from M_n that commute with all matrices from $\mathcal{S}$. When $\mathcal{S} = \{A\}$ we write shortly $\{A\}' = A'$. A matrix A is nonderogatory if its Jordan canonical form has exactly one Jordan block corresponding to each distinct eigenvalue.

Clearly, for $A \in M_n$ we have $A' = M_n$ if and only if A is a scalar matrix. In particular, $B' \subset (\lambda I)'$ for every $B \in M_n$ and every complex number λ. We will call a nonscalar matrix $A \in M_n$ maximal if every $B \in M_n$ satisfying $A' \subset B'$ and $A' \neq B'$ has to be a scalar matrix. The set of all nonscalar maximal matrices will be denoted by $\mathcal{M}$. Similarly, $A \in M_n$ is minimal if there is no $B \in M_n$ satisfying $B' \subset A'$ and $B' \neq A'$.

Lemma 3.1 *Let $A \in M_n$ be a nonscalar matrix. Then A is maximal if and only if either A is diagonalizable with exactly two eginvalues, or $A = \lambda I + N$ for some complex number λ and some square-zero matrix $N \neq 0$.*

Proof. Assume first that A is diagonalizable with exactly two eigenvalues and $B \in M_n$ a matrix satisfying $A' \subset B'$ and $A' \neq B'$. Then we may assume that

$$A = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$$

with $\lambda \neq \mu$. The commutant of A is the set of all matrices

$$\begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix}$$

where X and Y are any two square matrices of the appropriate size. It follows from $A' \subset B'$ that every such matrix commutes with B which further yields that

$$B = \begin{bmatrix} \delta & 0 \\ 0 & \tau \end{bmatrix}$$

for some complex numbers δ and τ. In the case $\delta \neq \tau$ we would have $A' = B'$, a contradiction. So, B has to be a scalar matrix, as desired.

Next, we will prove that also every matrix A of the form $A = \lambda I + N$ for some complex number λ and a square-zero matrix $N \neq 0$ is maximal. Replacing A by a similar matrix, if necessary, we may assume that

$$A = \begin{bmatrix} \lambda & I & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

for some complex number λ. The last column and the last row may be absent. Then the commutant of A is the set of all matrices of the form

$$\begin{bmatrix} X & Y & Z \\ 0 & X & 0 \\ 0 & U & V \end{bmatrix}$$

where X, Y, Z, U, V are arbitrary matrices of the appropriate size. Let $B \in M_n$ be a matrix satisfying $A' \subset B'$ and $A' \neq B'$. Similar argument as above yields that either

$$B = \begin{bmatrix} \mu & \delta I & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{bmatrix}$$

for some scalars μ, δ with $\delta \neq 0$, or $B = \mu I$. The first possibility cannot occur because $A' \neq B'$.

Now, if A has at least three eigenvalues, then it is similar to a matrix

$$\begin{bmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{bmatrix}$$

where A_1, A_2, and A_3 have pairwise disjoint spectra. We may assume that already the matrix A has this block diagonal form. It follows that the commutant of A is contained in the set of all matrices of the form

$$\begin{bmatrix} X & 0 & 0 \\ 0 & Y & 0 \\ 0 & 0 & Z \end{bmatrix}.$$

Then, obviosly, the matrix

$$B = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & -I \end{bmatrix}$$

is a nonscalar matrix whose commutant is larger than the commutant of A.

If A has two eigenvalues but is not diagonalizable then there is no loss of generality in assuming that

$$A = \begin{bmatrix} \lambda + M & 0 \\ 0 & \mu + N \end{bmatrix}$$

where $\lambda \neq \mu$ and M and N are nilpotents not both equal to zero. The nonscalar matrix

$$B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$$

has larger commutant than A, thus showing that A is not maximal also in this case.

The last case we have to treat is that A is of the form $A = \lambda I + N$ for some nilpotent with $N^2 \neq 0$. Using Jordan canonical form it is easy to verify that the commutant of A is a proper subset of the commutant of a nonscalar matrix $\lambda I + N^2$. This completes the proof.

Lemma 3.2 *Let $A \in M_n$. Then A is minimal if and only if A is nonderogatory.*

Proof. Assume first that A is nonderogatory and that $B \in M_n$ satisfies $B' \subset A'$. We have to show that $B' = A'$. From $B \in B'$ we conclude that A and B commute. It is well-known (and easy to verify) that if B commutes with a nonderogatory matrix A, then $B = p(A)$ for some polynomial p. It follows that $A' \subset B'$, as desired.

To prove the converse assume that A is in Jordan canonical form and that it has more than two Jordan blocks corresponding to the same eigenvalue λ. Denote these Jordan blocks by $J_1, \ldots, J_k$. Let B be a matrix obtained from A by replacing all diagonal entries in J_1 by μ_1, all diagonal entries in J_2 by μ_2,..., and all diagonal entries in J_k by μ_k, where $\mu_i \neq \mu_j$ whenever $i \neq j$. Then $B' \subset A'$ and $B' \neq A'$. This completes the proof.

Our next goal is to characterize matrices with n different eigenvalues using commutativity relations. Let A be a nonderogatory matrix. For two matrices $B, C \in A'$ the commutants B' and C' may be equal or different. We will take all matrices from A', then form the set of their commutants and denote by $\#A$ the cardinality of this set, $\#A = \text{card}\,\{B' : B \in A'\}$. The quantity $\#A$ does not change if we replace A by a similar matrix. So, we will assume that it is in the Jordan canonical form

$$A = \sum_{i=1}^{k} J_{n_i}(\lambda_i)$$

with $n_1 \geq \ldots \geq n_k$, $n_1 + \ldots + n_k = n$, and $\lambda_i \neq \lambda_j$ whenever $i \neq j$.

Assume first that $n_1 \geq 4$. Then, clearly, $B_\alpha = \alpha(E_{1,n_1-1} + E_{2,n_1}) + E_{1,n_1}$ belongs to A' and it is trivial to verify that $B'_\alpha \neq B'_\gamma$ whenever $\alpha \neq \gamma$. Hence, $\#A = \infty$ in this case.

The next case we will consider is that $n_2 \geq 2$. Then $B_\alpha = \alpha E_{1,n_1} + E_{n_1+1,n_1+n_2}$ belongs to A' and again $B'_\alpha \neq B'_\gamma$ whenever $\alpha \neq \gamma$. Thus, $\#A = \infty$ in this case as well.

Thus, we have proved the following result.

Lemma 3.3 *Let A be a nonderogatory matrix. If $\#A < \infty$, then either A has n different eigenvalues, or A has $n-1$ different eigenvalues, or A has $n-2$ different eigenvalues one of them being of algebraic multiplicity* 3.

In the next step we will consider only maximal matrices from the commutant A' of a nonderogatory matrix A. As before we form the set of their commutants and denote by $\#mA$ the cardinality of this set, $\#mA = \mathrm{card}\,\{B' : B \in A' \cap \mathcal{M}\}$. Assume first that A is diagonal with n different eigenvalues. Then every $B \in A' \cap \mathcal{M}$ is of the form $B = \alpha P + \beta(I-P)$ with $\alpha \neq \beta$ and P a diagonal idempotent, $P \neq 0, I$. Clearly, $B' = P'$. Two diagonal idempotents P and Q have the same commutant if and only if $P = Q$ or $P = I - Q$. Thus,

$$\#mA = \frac{1}{2}\left(\binom{n}{1} + \ldots + \binom{n}{n-1}\right) = 2^{n-1} - 1.$$

Now, let A be nonderogatory with $n-1$ different eigenvalues. Thus, its Jordan canonical form has one, say the first Jordan cell of the size 2×2, while all the others are 1×1 trivial Jordan cells. Hence, $B \in A' \cap \mathcal{M}$ if and only if $B = \alpha I + \beta E_{12}$ with $\beta \neq 0$ or B is diagonal with exactly two eigenvalues and the first two diagonal entries must be equal. Consequently, all the matrices $B \in A' \cap \mathcal{M}$ that are of the form scalar plus square-zero have the same commutant, and therefore,

$$\#mA = 1 + \frac{1}{2}\left(\binom{n-1}{1} + \ldots + \binom{n-1}{n-2}\right) = 2^{n-2}.$$

Similarly, if A has $n-2$ different eigenvalues one of them being of algebraic multiplicity 3, then $\#mA = 2^{n-3}$.

Hence, we have the following statement.

Lemma 3.4 *Let A be a nonderogatory matrix. If A has n different eigenvalues, then $\#mA = 2^{n-1} - 1$. If A has $n-1$ different eigenvalues, then $\#mA = 2^{n-2}$. If A has $n-2$ different eigenvalues one of them being of algebraic multiplicity* 3*, then $\#mA = 2^{n-3}$.*

Let n be any integer not smaller than 4. We denote $N_n = E_{12} + E_{23} + \ldots + E_{n-1,n}$.

Lemma 3.5 *Let $n \geq 4$. Assume that $A \in M_n$ commutes with N_n^2 and N_n^3. Then $A = a_0 I + a_1 N_n + a_2 N_n^2 + \ldots + a_{n-1} N_n^{n-1} + bE_{1,n-1} + cE_{2,n-1}$ for some scalars $a_0, \ldots, a_{n-1}, b, c$.*

Recall that N_n^2 is the matrix which has all entries on the second upper diagonal equal to 1 and all other entries equal to 0, N_n^3 is the matrix whose all entries on the third upper diagonal are equal to 1, while all other entries are equal to zero,... The conclusion of the above statement can be reformulated in the following way: Then $A = p(N_n) + R$ where p is a polynomial and R is a matrix whose all nonzero entries belong to the upper right 2×2 corner.

Proof. We denote by $e_1, \ldots, e_n$ the elements of the standard basis of the space of all $n \times 1$ matrices. Thus, e_i is the column matrix whose all entries are zero except the i-th entry which is equal to 1. Clearly, $E_{ij} = e_i e_j^t$, $1 \le i, j \le n$. Let $A = [a_{ij}]$ be a matrix that commutes with $N_n^2 = E_{13} + \ldots + E_{n-2,n}$ and $N_n^3 = E_{14} + \ldots + E_{n-3,n}$. Then

$$\sum_{j=1}^{n-2} E_{j,j+2} A = \sum_{j=1}^{n-2} A E_{j,j+2}$$

and

$$\sum_{j=1}^{n-3} E_{j,j+3} A = \sum_{j=1}^{n-3} A E_{j,j+3}.$$

Multiplying the first equation first by e_1 and then by e_2 on the right-hand side we get that the bottom left $(n-2) \times 2$ corner of A is zero. Now we multiply both equations by e_k^t on the left and by e_m on the right. We do this for all integers k, m satisfying $1 \le k \le n-2$ and $3 \le m \le n$ in the first case and all integers k, m satisfying $1 \le k \le n-3$ and $4 \le m \le n$ in the second case. We obtain

$$a_{k+2,m} = e_{k+2}^t A e_m = e_k^t \left(\sum_{j=1}^{n-2} e_j e_{j+2}^t A \right) e_m =$$

$$e_k^t \left(A \sum_{j=1}^{n-2} e_j e_{j+2}^t \right) e_m = e_k^t A e_{m-2} = a_{k,m-2}$$

for all k, m, $1 \le k \le n-2$, $3 \le m \le n$, and

$$a_{k+3,m} = a_{k,m-3}$$

for all k, m, $1 \le k \le n-3$, $4 \le m \le n$. The main diagonal of A has n entries, the first upper and the first lower diagonal have $n-1$ entries, the next two have $n-2$ entries,... The above two equations tell us that all the diagonals with at least 4 entries have all entries equal. Moreover, we have $a_{n-2,1} = a_{n,3}$ and $a_{1,n-2} = a_{3,n}$. Assume that $n > 4$. Applying the fact that the bottom left $(n-2) \times 2$ corner of A is zero we conclude that A is upper triangular matrix. Therefore, A is a sum of an upper triangular Toeplitz matrix and a matrix whose all nonzero entries belong to the upper right 2×2 corner. In the case $n = 4$ we know that A is of the form

$$\begin{bmatrix} a & b & * & * \\ c & a & * & * \\ 0 & 0 & a & b \\ 0 & 0 & c & a \end{bmatrix}.$$

In order to complete the proof we have to show that $c = 0$. This follows directly from the fact that A commutes with $N_4^3 = E_{14}$.

4 Maps on rank one idempotents

This section will be devoted to the proofs of the two theorems on orthogonality preserving maps on rank one idempotents.

Proof of Theorem 2.3. Recall that every idempotent P of rank one can be written as $P = xy^t$ where x and y are $n \times 1$ matrices satisfying $y^t x = 1$. The space of all $n \times 1$ matrices will be identified with $\mathbb{C}^n$. For two idempotents of rank one $P = xy^t$ and $Q = uv^t$ we write $P \sharp Q$ if x and u are linearly dependent or y and v are linearly dependent. Our first step will be to show for $P, Q \in I_n$ we have $P \sharp Q$ if and only if there exist orthogonal sets $\{S, R_4, \dots, R_n\} \subset \{P, Q\}^\perp$ and $\{T, R_4, \dots, R_n\} \subset \{P, Q\}^\perp$ with $S \neq T$. To see this assume first that $xy^t = P \sharp Q = uv^t$. If $P = Q$ then we can find an orthogonal set $\{R_1, \dots, R_n\} \subset I_n$ with $P = Q = R_1$. Choosing $S = R_2$ and $T = R_3$ we get the orthogonal sets of rank one idempotents with the desired properties. So, let us assume that $P \neq Q$. We have that either x and u are linearly dependent, or y and v are linearly dependent. We will consider only the first possibility. After replacing P and Q by simultaneously similar matrices, if necessary, we may assume that $P = E_{11}$ and $Q = E_{11} + E_{12}$. Set $R_k = E_{kk}$, $k = 4, \dots, n$, $S = E_{33}$, and $T = E_{32} + E_{33}$. It is then easy to verify that $\{S, R_4, \dots, R_n\}$ and $\{T, R_4, \dots, R_n\}$ are orthogonal subsets of the set $\{P, Q\}^\perp$. In the case that $P \not\sharp Q$ both pairs of vectors x, u and y, v are linearly independent. Let $\{S, R_4, \dots, R_n\} \subset \{P, Q\}^\perp$ and $\{T, R_4, \dots, R_n\} \subset \{P, Q\}^\perp$ be orthogonal sets of rank one idempotents with $R_k = z_k w_k^t$, $k = 4, \dots n$. Let us prove that $x, u, z_4, \dots, z_n$ are linearly independent vectors. Let $\lambda x + \delta u + \sum_{k=4}^n \eta_k z_k = 0$. Because $w_4^t x = w_4^t u = w_4^t z_5 = \dots = w_4^t z_n = 0$ we have $\eta_4 = 0$. Similarly, all other η's must be zero, and because of linear independence of x and u, the scalars λ and δ have to be zero as well. Similarly, vectors $y, v, w_4, \dots, w_n$ are linearly independent. Now, both S and T are orthogonal to P, Q, and $R_4, \dots, R_n$. Therefore, $y^t S = v^t S = w_4^t S = \dots = w_n^t S = 0$, and consequently, the column space of S is equal to the one-dimensional space $\{a \in \mathbb{C}^n : y^t a = v^t a = w_4^t a = \dots = w_n^t a = 0\}$. The same is true for the column space of T. Similarly we prove that S and T have the same row spaces. In other words, S is a scalar multiple of T. But they are both idempotents, and therefore, $S = T$, as desired.

Assume now that $P \sharp Q$. Then, by the previous step, we can find orthogonal sets $\{S, R_4, \dots, R_n\} \subset \{P, Q\}^\perp$ and $\{T, R_4, \dots, R_n\} \subset \{P, Q\}^\perp$ with $S \neq T$. Because ϕ is injective and preserves orthogonality, the ξ-images of these idempotents have the same properties, and applying again the characterization of the relation $\sharp$ we conclude that $\xi(P) \sharp \xi(Q)$.

For every nonzero $x \in \mathbb{C}^n$ we set $L_x = \{xu^t : u \in \mathbb{C}^n \text{ and } u^t x = 1\} \subset I_n$.

Similarly, for every nonzero $y \in \mathbb{C}^n$ we define $R_y = \{vy^t : v \in \mathbb{C}^n \text{ and } y^t v = 1\} \subset I_n$. Clearly, if $xu^t \neq xw^t$ both belong to L_x, then u and w are linearly independent. For every nonzero x we will call L_x a set of rank one idempotents of type I and R_x a set of rank one idempotents of type II.

Our next goal is to show that every set of rank one idempotents of type I is mapped either into a set of rank one idempotents of type I, or into a set of rank one idempotents of type II. Indeed, let x be any nonzero vector and let xu_1^t and xu_2^t be two different elements of L_x. By the previous step, either

$$\xi(xu_1^t) = zv_1^t \quad \text{and} \quad \xi(xu_2^t) = zv_2^t$$

for some vectors z, v_1, and v_2, or

$$\xi(xu_1^t) = w_1 y^t \quad \text{and} \quad \xi(xu_2^t) = w_2 y^t$$

for some vectors y, w_1, and w_2. Let us consider just the first case as the proof in the second case is almost the same. For an arbitrary vector u_3 satisfying $u_3^t x = 1$ we have $\xi(xu_3^t) \sharp zv_1^t$ and $\xi(xu_3^t) \sharp zv_2^t$. But v_1 and v_2 are linearly independent. Therefore, $\xi(xu_3^t) \in L_z$. Hence, $\xi(L_x) \subset L_z$.

Clearly, the same statement holds true also for sets of rank one idempotents of type II, that is, every set of rank one idempotents of type II is mapped either into a set of rank one idempotents of type I, or into a set of rank one idempotents of type II.

After composing ξ by the transposition, if necessary, we may assume that there is a set of rank one idempotents of type I that is mapped into a set of rank one idempotents of type I. We will prove that then every set of rank one idempotents of type I is mapped into a set of rank one idempotents of type I. We first observe that if $x, y \in \mathbb{C}^n$ are linearly independent then we can find $P_1, P_2 \in L_x$ and $Q_1, Q_2 \in L_y$ such that $P_1 \neq P_2$, $Q_1 \neq Q_2$, and $P_i \sharp Q_i$, $i = 1, 2$. As x and y are linearly independent we have also $P_i \neq Q_j$ for all pairs $i, j \in \{1, 2\}$. Further, we claim that for any pair of nonzero vectors $x, y \in \mathbb{C}^n$ the relations $P_1, P_2 \in L_x$, $Q_1, Q_2 \in R_y$ and $P_i \sharp Q_i$, $i = 1, 2$, imply that $y^t x \neq 0$ and either $P_1 = P_2$, or $Q_1 = Q_2$, or $P_i = Q_j$ for some $i, j = 1, 2$. Indeed, after applying a similarity, we may and we do assume that $x = e_1$. If $y^t x = 0$ we may assume that $x = e_1$ and $y = e_2$. It is then clear that $P \in L_x$ and $Q \in R_y$ imply $P \not\sharp Q$. So, $y^t x \neq 0$ and we may assume without loss of generality (after applying a similarity and multiplying y by a nonzero scalar) that $x = y = e_1$. From $P_1 \sharp Q_1$ and $P_2 \sharp Q_2$ we get now immediately that at least two of the idempotents P_1, P_2, Q_1, Q_2 are equal to E_{11}, as desired. Using the injectivity assumption and the implication $P \sharp Q \Rightarrow \xi(P) \sharp \xi(Q)$ we can now easily conclude that every set of rank one idempotents of type I is mapped into a set of rank one idempotents of type I.

For a nonzero $x \in \mathbb{C}^n$ we denote by $[x]$ the one-dimensional space spanned by x. As usual, $\mathbb{PC}^n = \{[x] : x \in \mathbb{C}^n \setminus 0\}$. We have proved that for every nonzero vector x there is a nonzero vector u such that $\xi(L_x) \subset L_u$. Thus, ξ

induces a map η on $\mathbb{PC}^n$ such that $[u] = \eta([x])$ if and only if $\xi(L_x) \subset L_u$. Assume that $[x] \subset [u] + [v]$ for some nonzero $x, u, v \in \mathbb{C}^n$. We want to prove that $\eta([x]) \subset \eta([u]) + \eta([v])$. There is nothing to prove if u and v are linearly dependent. So, assume that they are linearly independent. Then we can find a maximal orthogonal set of rank one idempotents $\{P_1, \ldots, P_n\}$ such that the column space of P_1 is $[u]$ and the column space of P_2 is $[v]$. By the orthogonality preserving property we have $\xi(P_k)\eta([u]) = \xi(P_k)\eta([v]) = \{0\}$, $k = 3, \ldots, n$. Since $\xi(P_1), \ldots, \xi(P_n)$ are orthogonal we have $\eta([u]) \neq \eta([v])$ and every vector w satisfying $\xi(P_k)w = 0$ for all $k = 3, \ldots, n$, belongs to the direct sum $\eta([u]) \oplus \eta([v])$. We can find a rank one idempotent $R \in L_x$ that is orthogonal to all $P_3, \ldots, P_n$. So, the above direct sum contains the column space of $\xi(R)$. In other words, we have $\eta([x]) \subset \eta([u]) + \eta([v])$, as desired. Hence, we can apply the nonsurjective version of the fundamental theorem of projective geometry [6, Theorem 3.1] to conclude that there exists an endomorphism f of the complex field and a linear map $T : \mathbb{C}^n \to \mathbb{C}^n$ such that $\xi(L_x) \subset L_u$, where $u = Tx^f$. Here,

$$x^f = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}^f = \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix}.$$

If $w_1, \ldots, w_n \in \mathbb{C}^n$ are linearly independent, then we can find $P_i \in L_{w_i}$, $i = 1, \ldots, n$, such that $P_i \perp P_j$ whenever $i \neq j$. Thus, $\xi(P_i) \perp \xi(P_j)$ whenever $i \neq j$, and consequently, $Tw_1^f, \ldots, Tw_n^f$ are linearly independent. It follows that T is invertible and after replacing ξ by the map $P \mapsto T^{-1}\xi(P)T$, we may assume that $\xi(L_x) \subset L_{x^f}$. In other words, for every $xy^t \in I_n$ there exists $u \in \mathbb{C}^n$ such that $\xi(xy^t) = x^f u^t$. Assume that a nonzero $w \in \mathbb{C}^n$ satisfies $y^t w = 0$. Then, since $y^t x = 1$, the vectors x and w are linearly independent. So we can find a vector $z \in \mathbb{C}^n$ such that $z^t x = 0$ and $z^t w = 1$. It follows that $xy^t \perp wz^t$, and consequently, $u^t w^f = 0$. Since this holds for every vector w with $y^t w = 0$, the vector u has to be a scalar multiple of y^f. Now, $u^t x^f = 1$, and therefore, $u = y^f$. Hence, $\xi(xy^t) = x^f (y^f)^t$. This completes the proof.

Proof of Theorem 2.4. Let $x \otimes f, y \otimes g \in I(X)$. Similarly as in the finite-dimensional case we define that $x \otimes f \,\sharp\, y \otimes g$ if x and y are linearly dependent or f and g are linearly dependent. Let us start with some simple observations. Assume that $P = x \otimes f, Q = x \otimes g \in I(X)$. A rank one idempotent $T \in I(X)$ belongs to $\{P, Q\}^\perp$ if and only if $TP = PT = TQ = QT = 0$, or equivalently, $Tx = 0$ and $T'f = T'g = 0$. From here we will conclude that $R \in I(X)$ belongs to $\left(\{P, Q\}^\perp\right)^\perp$ if and only if $R = x \otimes (\lambda f + (1 - \lambda)g)$ for some $\lambda \in \mathbb{C}$. Indeed, if $R = u \otimes h \in \left(\{P, Q\}^\perp\right)^\perp$, then $RT = TR = 0$ for every $T \in I(X)$ satisfying $Tx = 0$ and $T'f = T'g = 0$. We first have to prove that u and x are linearly dependent. Assume to the contrary that this is not true. Then we can find a vector $z \in X$ such that the set $\{x, z, u\}$ is linearly independent and

$f(z) = g(z) = 0$. Then there exists $k \in X'$ with $k(x) = 0$ and $k(z) = k(u) = 1$. Hence, $T = z \otimes k$ satisfies $Tx = 0$ and $T'f = T'g = 0$ but also $TR = z \otimes h \neq 0$, a contradiction. This contradiction shows that we may assume, after absorbing a constant, if necessary, that $u = x$. If h does not belong to the linear span of f and g, then we can find a vector z such that $f(z) = g(z) = 0$ and $h(z) = 1$. Because $f(x) = 1$, the vectors z and x are linearly independent and therefore there is a functional $k \in X'$ such that $k(z) = 1$ and $k(x) = 0$. Then once again $T = z \otimes k \in I(X)$ satisfies $Tx = 0$ and $T'f = T'g = 0$ but also $RT = x \otimes k \neq 0$, a contradiction. Thus, $R = x \otimes (\alpha f + \beta g)$ for some scalars α, β. It follows from $(\alpha f + \beta g)(x) = 1$ that $\alpha + \beta = 1$, as desired. To prove the converse we have to show that every rank one idempotent $R = x \otimes (\lambda f + (1 - \lambda)g)$ belongs to $\left(\{P, Q\}^\perp\right)^\perp$. So, we have to see that $TR = RT = 0$ for every $T \in I(X)$ with $Tx = 0$ and $T'f = T'g = 0$. This is obviously true.

We are now ready to show that for $P, Q \in I(X)$, $P \neq Q$, we have $P \sharp Q$ if and only if for every pair $R, S \in \left(\{P, Q\}^\perp\right)^\perp$ with $R \neq S$ we have $\{P, Q\}^\perp = \{R, S\}^\perp$. Assume first that $P = x \otimes f \sharp y \otimes g = Q$. Then either x and y are linearly dependent, or f and g are linearly dependent. We will consider only the first possibility. So, we may assume that $x = y$. Let $R, S \in \left(\{P, Q\}^\perp\right)^\perp$ with $R \neq S$. Then by the previous step we have $R = x \otimes (\lambda f + (1 - \lambda)g)$ and $S = x \otimes (\mu f + (1 - \mu)g)$ for some scalars $\lambda \neq \mu$. It is now straightforward to check that $\{P, Q\}^\perp = \{R, S\}^\perp$.

Assume now that $P = x \otimes f \not\sharp y \otimes g = Q$. Then both pairs x, y and f, g are linearly independent. Choose $R = P = x \otimes f$ and $S = x \otimes (g + (1 - g(x))f)$. It is straightforward to see that $T \in I(X)$ belongs to $\{P, Q\}^\perp$ if and only if $Tx = Ty = 0$ and $T'f = T'g = 0$. It follows easily that $R, S \in \left(\{P, Q\}^\perp\right)^\perp$. Clearly, $R \neq S$. We can find a vector z satisfying $f(z) = g(z) = 0$ such that x, z, y are linearly independent. Then there exists $k \in X'$ such that $k(z) = k(y) = 1$ and $k(x) = 0$. Hence, $z \otimes k$ is an idempotent belonging to $\{R, S\}^\perp$ but $z \otimes k \cdot Q \neq 0$. Thus, $\{P, Q\}^\perp \neq \{R, S\}^\perp$, as desired.

For a nonzero $x \in X$ and a nonzero $f \in X'$ we denote $L_x = \{x \otimes g \,:\, g \in X' \text{ and } g(x) = 1\} \subset I(X)$ and $R_f = \{y \otimes f \,:\, y \in X \text{ and } f(y) = 1\} \subset I(X)$. As in the finite-dimensional case we can prove that either for every nonzero $x \in X$ there exists a nonzero $y \in X$ such that $\xi(L_x) \subset L_y$, or for every nonzero $x \in X$ there exists a nonzero $f \in X'$ such that $\xi(L_x) \subset R_f$. In fact, since ξ is bijective and preserves orthogonality in both directions we have $\xi(L_x) = L_y$ in the first case and $\xi(L_x) = R_f$ in the second case. We will consider only the second case which requires slightly more complicated arguments than the first one. So, ξ induces a bijective map $\eta : \mathbb{P}X \to \mathbb{P}X'$ such that $\eta([x]) = [f]$ if and only if $\xi(L_x) = R_f$. Assume that $[x] \not\subset [y] + [z]$ for some nonzero vectors x, y, z. We want to prove that $\eta([x]) \not\subset \eta([y]) + \eta([z])$. We may assume that y and z are linearly independent since otherwise there is nothing to prove. Choose $f, g, h \in X'$ such that $f(x) = g(y) = h(z) = 1$ and $f(y) = f(z) = g(x) = g(z) = h(x) = h(y) = 0$. Then $\xi(x \otimes f)$, $\xi(y \otimes g)$, and $\xi(z \otimes h)$ are pairwise orthogonal which implies that the linear span of $\eta([x]), \eta([y]), \eta([z])$ is of dimension 3. This yields the desired relation $\eta([x]) \not\subset \eta([y])+\eta([z])$. We can prove the same for the inverse of η. Thus, for any nonzero $x, y, z \in X$ we have $[x] \subset [y]+[z]$ if and only if $\eta([x]) \subset \eta([y]) + \eta([z])$. By the fundamental theorem of projective geometry there exists a bijective semilinear map $S : X \to X'$ such that $\eta([x]) = [Sx]$. We claim that S carries closed hyperplanes of X to closed hyperplanes of X'. Let $W \subset X$ be a closed hyperplane. Choose $x \in X \setminus W$. We define $f \in X'$ by $f(x) = 1$ and $f(W) = \{0\}$. We have $\xi(x \otimes f) = u \otimes Sx$ for some $u \in X$. All we have to do is to show that $SW = \{g \in X' : g(u) = 0\}$. Let $y \in W$ be a nonzero vector. We can find $k \in X'$ with $k(y) = 1$ and $k(x) = 0$. Then $x \otimes f \perp y \otimes k$. We have $\xi(y \otimes k) = w \otimes Sy$ for some $w \in X$. So, $(w \otimes Sy)(u \otimes Sx) = 0$, and consequently, $Sy \in \{g \in X' : g(u) = 0\}$. Thus, $SW \subset \{g \in X' : g(u) = 0\}$ and because both subspaces are of codimension one, they have to be equal. We prove similarly that the inverse of S carries closed hyperplanes to closed hyperplanes. It follows from [7, Lemma 3] that S is continuous and linear or conjugate linear.

We also know that either for every $f \in X'$ there exists $x \in X$ such that $\phi(R_f) = L_x$, or for every $f \in X'$ there exists $g \in X'$ such that $\phi(R_f) = R_g$. The second case cannot occur because each R_g is a ξ-image of some L_u. Now, using the same approach as above we conclude that there exists a bounded bijective linear or conjugate linear map $T : X' \to X$ such that for every $x \otimes f \in I(X)$ we have $\xi(x \otimes f) = Tf \otimes g$ for some $g \in X'$. Thus, for every $x \otimes f \in I(X)$ we have $\xi(x \otimes f) = \frac{1}{(Sx)(Tf)} Tf \otimes Sx$. In particular, $f(x) = 1$ yields that $(Sx)(Tf) \neq 0$. Because S and T are semilinear we have for every pair $x \in X$ and $f \in X'$ the implication $f(x) \neq 0 \Rightarrow (Sx)(Tf) \neq 0$.

Now, let $x \in X$ and $f \in X'$ satisfy $f(x) = 0$. Find $g \in X'$ such that $g(x) = 1$. Then $(g + \lambda f)(x) = 1$ for every complex number λ, and therefore, $(Sx)(Tg + \mu Tf) \neq 0$ for every $\mu \in \mathbb{C}$. This is possible only if $(Sx)(Tf) = 0$.

Hence, we have $(Sx)(Tf) = 0$ if and only if $f(x) = 0$. Here we have to distinguish two cases. We will consider only the case that S is conjugate linear since the linear case goes through in the same way. We claim that then T is conjugate linear as well and there exists a complex constant c such that $(Sx)(Tf) = c\,\overline{f(x)}$. Indeed, choose any $x \otimes f \in I(X)$ and set $c = (Sx)(Tf)$. Consider $u \in X$ and $g \in X'$ such that $g(u) = 1$ and $g(x) = f(u) = 0$. Then $(f + g)(x - u) = 0$ which yields $(Su)(Tg) = (Sx)(Tf) = c$. Let now $w \otimes h$ be any member of $I(X)$ and we want to show that $(Sw)(Th) = c$. For this purpose we choose $z \in X$ such that $f(z) = h(z) = 0$ and $z \notin \operatorname{span}\{x, w\}$. Choose also $k \in X'$ satisfying $k(w) = k(x) = 0$ and $k(z) = 1$. As before we prove that $(Sz)(Tk) = c$ and $(Sz)(Tk) = (Sw)(Th)$ which implies the desired relation $(Sw)(Th) = c$ for every pair $w \in X, h \in X'$ with $h(w) = 1$. This further implies that also T is conjugate linear, and consequently, we have $(Sw)(Th) = c\overline{h(w)}$ for every pair $w \in X$, $h \in X'$.

Replacing S by $c^{-1}S$ we may assume that $(Sx)(Tf) = \overline{f(x)}$ for every pair $x \in X$, $f \in X'$. Recall that if Y, W are Banach spaces and if $A : Y \to W$

is a bounded conjugate linear operator, then $A' : W' \to Y'$ is defined by $(A'k)(z) = \overline{k(Az)}$, $z \in Y$, $k \in W'$. Let K be the natural embedding of $X \to X''$. Then $S = (T^{-1})'K$. Because both S and T are bijective, the embedding K is also bijective and $\xi(x \otimes f) = T(f \otimes Kx)T^{-1}$, $x \otimes f \in I(X)$. This completes the proof.

5 Commutativity preserving maps

Now we are ready to start our study of commutativity preserving maps. We will first treat such maps without imposing the continuity assumption. So, the goal of this section is to prove Theorem 2.1. Thus, let us assume that $n \geq 3$ and that $\phi : M_n \to M_n$ is a bijective map preserving commutativity in both directions. Then, obviously, for every subset $\mathcal{S} \subset M_n$ we have $\phi(\mathcal{S}') = \phi(\mathcal{S})'$. If $A \in M_n$ has n different eigenvalues, then A is diagonalizable and every $A \in B'$ is simultaneously diagonalizable. Assume that such an A is already in a diagonal form and that $B \in A'$. Then the commutant B' is completely determined if we know which of the diagonal entries of B are equal. Thus, $\sharp A < \infty$. It follows from Lemmas 3.1, 3.2, 3.3, and 3.4 that ϕ maps the set of all matrices with n different eigenvalues onto itself. Further, a matrix A is diagonalizable if and only if it commutes with some matrix with n different eigenvalues. Thus, $\mathcal{D}$, the set of all diagonalizable matrices is mapped by ϕ onto itself. Denote by $\mathcal{D}_k$, $k = 1, \ldots, n$, the set of all diagonalizable matrices with exactly k eigenvalues. We have $A \in \mathcal{D}_1$ if and only if $A = \lambda I$ for some $\lambda \in \mathbb{C}$ and this is equivalent to $A' = M_n$. Thus $\mathcal{D}_1$ is mapped onto itself. The same is true for $\mathcal{D}_2 = \mathcal{M} \cap \mathcal{D}$. Observe that for $A \in \mathcal{D}$ the following two statements are equivalent:

- $A \in \mathcal{D}_3$,
- $A \notin \mathcal{D}_1 \cup \mathcal{D}_2$ and every matrix $B \in \mathcal{D}$ satisfying $B \in A'$, $A' \subset B'$, and $A' \neq B'$ belongs to $\mathcal{D}_1 \cup \mathcal{D}_2$.

It follows easily that $\phi(\mathcal{D}_3) = \mathcal{D}_3$. Repeating this procedure we get $\phi(\mathcal{D}_k) = \mathcal{D}_k$, $k = 1, \ldots, n$.

We denote by $\mathcal{Q} \subset \mathcal{D}_2$ the set of all matrices of the form $\lambda P + \mu(I - P)$, where $\lambda \neq \mu$ and P is an idempotent of rank one. So, $\mathcal{Q}$ is the set of all diagonalizable matrices with exactly two eigenvalues one of them having the eigenspace of dimension one. In our next step we will prove that ϕ maps the set $\mathcal{Q}$ onto itslef. In the case $n = 3$ we have $\mathcal{Q} = \mathcal{D}_2$ and so, there is nothing to prove. Therefore we will assume in this paragraph that $n \geq 4$. We will verify that for $A \in \mathcal{D}_2$ the following two statements are equivalent:

- $A \in \mathcal{Q}$,
- for every $B \in A' \cap \mathcal{D}_2$ we have $\{A, B\}'' \subset \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$.

Assume for a moment that we have already proved this. Then, because ϕ preserves the first commutants, it has to preserve also the second commutants and since it preserves $\mathcal{D}_k$, $k = 1, 2, 3$, we have necessarily $\phi(\mathcal{Q}) = \mathcal{Q}$, as desired. So, assume that $A = \lambda P + \mu(I - P) \in \mathcal{Q}$ and $B \in A' \cap \mathcal{D}_2$. A matrix C commutes with A if and only if it commutes with P. So, there is no loss of generality in assuming that already A is an idempotent of rank one, and after applying a similarity, if necessary, we may assume that $A = E_{11}$. Moreover, two diagonalizable matrices commute if and only if they are simultaneously diagonalizable, and therefore, there is no loss of generality in assuming that $B = \tau(E_{11} + \ldots + E_{kk}) + \delta(E_{k+1,k+1} + \ldots + E_{nn})$, $1 \le k \le n-1$, $\tau \ne \delta$. If $k = 1$, then $\{A, B\}'' = \mathrm{span}\,\{E_{11}, I - E_{11}\} \subset \mathcal{D}_1 \cup \mathcal{D}_2$, and if $2 \le k \le n-1$, then $\{A, B\}'' = \mathrm{span}\,\{E_{11}, E_{22} + \ldots + E_{kk}, I - (E_{11} + \ldots + E_{kk})\} \subset \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$. To prove the other direction assume that $A \in \mathcal{D}_2 \setminus \mathcal{Q}$. As before there is no loss of generality in assuming that $A = E_{11} + \ldots + E_{kk}$ for some k, $2 \le k \le n-2$. Take $B = E_{11} + E_{k+1,k+1}$ and observe that then $\{A, B\}'' = \mathrm{span}\,\{E_{11}, E_{22} + \ldots + E_{kk}, E_{k+1,k+1}, I - (E_{11} + \ldots + E_{k+1,k+1})\}$ contains matrices with four different eigenvalues.

To each $A \in \mathcal{Q}$ we associate the unique idempotent P of rank one satisfying $A = \lambda P + \mu(I - P)$, $\lambda, \mu \in \mathbb{C}$. If $A, B \in \mathcal{Q}$ and P and Q are the corresponding idempotents of rank one, then $P = Q$ if and only if $A' = B'$. Thus, ϕ induces a bijective map $\xi : I_n \to I_n$. Moreover, we have $P \perp Q$ if and only if A and B commute and $A' \ne B'$. Thus, the map ξ preserves the orthogonality in both directions. By Theorem 2.3, there exists a nonsingular matrix $T \in M_n$ and an automorphism $f : \mathbb{C} \to \mathbb{C}$ such that either $\xi(P) = TP_fT^{-1}$, $P \in I_n$, or $\xi(P) = TP_f^tT^{-1}$, $P \in I_n$. Replacing ϕ by $A \mapsto T^{-1}\phi(A_{f^{-1}})T$, and composing the obtained map with the transposition, if necessary, we may assume without loss of generality that for every idempotent P of rank one the set of all matrices of the form $\lambda P + \mu(I - P)$, $\lambda \ne \mu$, is mapped bijectively onto itself. In other words, for every $A \in \mathcal{Q} \cup \mathbb{C}I$ there exist polynomials p_A and q_A such that $\phi(A) = p_A(A)$ and $A = q_A(p_A(A))$. Hence, after composing ϕ by an appropriate regular locally polynomial map (this map acts like the identity outside $\mathcal{Q} \cup \mathbb{C}I$), we may assume that $\phi(A) = A$ for every $A \in \mathcal{Q} \cup \mathbb{C}I$.

In the next step we will prove that after composing ϕ by yet another regular locally polynomial map we may assume that $\phi(A) = A$ for every diagonalizable A. As before, we need to show that for every diagonalizable A there are polynomials p_A and q_A such that $\phi(A) = p_A(A)$ and $A = q_A(p_A(A))$. In fact, it is enough to prove this only for diagonal matrices. Indeed, assume that we have proved the existenece of such polynomials for diagonal matrices and let A be any diagonalizable matrix. Then there is an invertible $R \in M_n$ such that $RAR^{-1} = D$ is diagonal. The map $\psi(X) = R\phi(R^{-1}XR)R^{-1}$ is a bijective map preserving commutativity in both directions with the additional property that $\psi(A) = A$ for every $A \in \mathcal{Q} \cup \mathbb{C}I$. Thus, by our assumption, $\psi(D)$ and D have the same commutant, or equivalently, $\phi(A)$ and A have the same commutant which is the same as the existence of polynomials p_A and q_A such that $\phi(A) = p_A(A)$

and $A = q_A(p_A(A))$.

Hence, let D be a diagonal matrix. It is easy to see that $D' = \mathrm{span}\,(I_n \cap D')$. Since ϕ acts like the identity on I_n we have $\phi(D)' = D'$, as desired. Thus, from now on we will assume that $\phi(A) = A$ for every diagonalizable matrix A.

Let $\mathcal{F}$ be the union of the set of all diagonalizable matrices and the set of all matrices that can be written as $\lambda I + N$ where λ is any complex number and N is any nilpotent matrix of rank one. We will prove in this paragraph that after composing ϕ by an appropriate regular locally polynomial map we may assume that $\phi(A) = A$ for every $A \in \mathcal{F}$. As before, it is enough to show that $\phi(N)' = N'$ for every nilpotent of rank one. And to do this we have to verify this equality only for the special case when $N = E_{12}$. Since diagonalizable matrices are mapped identically onto diagonalizable matrices we get from Lemma 3.1 that $\phi(E_{12})$ is a scalar plus a nonzero square-zero matrix. Further we know that $\phi(E_{12})$ commutes with $E_{11} + E_{22}$, $E_{33}, \ldots, E_{nn}$. All these yield that $\phi(E_{12})$ is a scalar plus a nonzero square-zero matrix M, where M has nonzero entries only in the upper left 2×2 corner. Now we apply the fact that E_{12} commutes with a rank two idempotent $E_{11} + E_{22} + E_{13}$ to conclude that the first column of M has to be zero. Since M is nilpotent, it has to be a scalar multiple of E_{12}. This completes the proof of this step.

Now we are ready to complete the proof. We know that $\phi(A) = A$ for every diagonalizable matrix A and every A that is a sum of a scalar matrix and a rank one nilpotent. We want to prove that for every $A \in \mathcal{C}$ there is a polynomial p_A such that $\phi(A) = p_A(A)$ and $\phi(A)' = (p_A(A))'$.

For a pair of complex numbers λ, a we denote by $J(\lambda, a)$ the 2×2 matrix

$$J(\lambda, a) = \begin{bmatrix} \lambda & a \\ 0 & \lambda \end{bmatrix}.$$

Let S be an arbitrary $n \times n$ invertible matrix, k, m nonnegative integers with $2k + m = n$, and $\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_m$ complex numbers. We have to prove that

$$A = S\mathrm{diag}\,(J(\lambda_1, 1), \ldots, J(\lambda_k, 1), \mu_1, \ldots, \mu_m) S^{-1}$$

is mapped into

$$\phi(A) = S\mathrm{diag}\,(J(\tau_1, a_1), \ldots, J(\tau_k, a_k), \xi_1, \ldots, \xi_m) S^{-1},$$

where $\lambda_i = \lambda_j$ if and ony if $\tau_i = \tau_j$, $\mu_i = \mu_j$ if and ony if $\xi_i = \xi_j$, $\lambda_i = \mu_j$ if and ony if $\tau_i = \xi_j$, $a_i = a_j$ whenever $\lambda_i = \lambda_j$, and $a_i \neq 0$ for every $i = 1, \ldots, k$.

Because A commutes with idempotents

$$S\mathrm{diag}\,(I, \ldots, 0, 0, \ldots, 0) S^{-1},$$

$$\vdots$$

$$S\mathrm{diag}\,(0, \ldots, I, 0, \ldots, 0) S^{-1},$$

$$S\mathrm{diag}\,(0,\ldots,0,1,\ldots,0)S^{-1},$$
$$\vdots$$
$$S\mathrm{diag}\,(0,\ldots,0,0,\ldots,1)S^{-1},$$

the matrix $\phi(A)$ commutes with these idempotents as well, and therefore,

$$\phi(A) = S\mathrm{diag}\,(A_1,\ldots,A_k,\xi_1,\ldots,\xi_m)S^{-1},$$

where $A_1,\ldots,A_k$ are 2×2 matrices, and $\xi_1,\ldots,\xi_m \in \mathbb{C}$. The matrix A commutes with

$$S(J(0,1)\oplus 0)S^{-1} = SE_{12}S^{-1} \in \mathcal{F}$$

and, of course, the same must be true for $\phi(A)$. Thus, $A_1 = J(\tau_1, a_1)$ for some complex numbers τ_1, a_1. If $a_1 = 0$, then $\phi(A)$ commutes with every matrix

$$S\mathrm{diag}\,(P,\ldots,0,0,\ldots,0)S^{-1},$$

where P is any 2×2 idempotent of rank one and then the same must be true for A. This contradiction shows that $a_1 \neq 0$. Similarly, we see that all the matrices A_i have a similar form. So, we have proved that

$$\phi(A) = S\mathrm{diag}\,(J(\tau_1,a_1),\ldots,J(\tau_k,a_k),\xi_1,\ldots,\xi_m)S^{-1},$$

for some complex numbers $\tau_1,\ldots,\tau_k$, $\xi_1,\ldots,\xi_m$, and some nonzero complex numbers $a_1,\ldots,a_k$. Assume that two of the λ's, say λ_1 and λ_2, are equal. Then A commutes with $SE_{14}S^{-1} \in \mathcal{F}$. It follows that $\phi(A)$ commutes with $SE_{14}S^{-1}$ which further yields that $\tau_1 = \tau_2$. The same argument shows that $\tau_1 = \tau_2$ implies that $\lambda_1 = \lambda_2$. Hence, $\lambda_i = \lambda_j$ if and ony if $\tau_i = \tau_j$, and similarly, $\mu_i = \mu_j$ if and only if $\xi_i = \xi_j$.

If one of the λ's is equal to some μ, say $\lambda_1 = \mu_1$, then A commutes with $SE_{1,2k+1}S^{-1}$, which implies that $\tau_1 = \xi_1$. Similarly, if $\tau_i = \xi_j$ for some integers i,j, then $\lambda_i = \mu_j$.

It remains to prove that $\lambda_i = \lambda_j$ yields that $a_i = a_j$. Assume with no loss of generality that $\lambda_1 = \lambda_2$. Then A commutes with the diagonalizable matrix

$$D = S(\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \oplus 0)S^{-1}.$$

Here, I stands for the 2×2 identity matrix, and the last 0 denotes the $(n-4)\times(n-4)$ zero matrix. The matrix D has to commute with $\phi(A)$ as well. The desired equation $a_1 = a_2$ follows easily. This completes the proof.

6 Example

We define $\mathcal{N} \subset M_n$ to be the set of all matrices of the form $cI + N$ where c is any complex number and N is a nilpotent of maximal nilindex, $N^n = 0$ and $N^{n-1} \neq 0$. Obviously, for $A = cI + N \in \mathcal{N}$ the scalar c and the nilpotent N are uniquely determined. Using Jordan canonical form we easily see that for $A = cI + N \in \mathcal{N}$ we have $A' = N' = \{p(N) : p \in \mathcal{P}\}$. Of course, $B = \sum_{k=0}^{n-1} \lambda_k N^k \in A'$ belongs to $\mathcal{N}$ if and only if $\lambda_1 \neq 0$. If $A, B \in \mathcal{N}$ we will write $A \sim B$ if $A' = B'$. Clearly, for $A, B \in \mathcal{N}$ we have $A \sim B$ if and only if $AB = BA$ and this is further equivalent to $A = p(B)$ and $B = q(A)$ for some $p, q \in \mathcal{P}$. Further, for $A = cI + N \in \mathcal{N}$ and $B = dI + M \in \mathcal{N}$ we write $A \approx B$ if $A' \setminus \mathcal{N} = \text{span}\,\{I, N^2, N^3, \ldots, N^{n-1}\} = \text{span}\,\{I, M^2, M^3, \ldots, M^{n-1}\} = B' \setminus \mathcal{N}$. Clearly, $A \sim B$ yields that $A \approx B$. Thus, the relation $\approx$ induces an equivalence relation on $\mathcal{N}/_\sim = \{[A] : A \in \mathcal{N}\}$, the set of all equivalence classes with respect to $\sim$. If $[A], [B] \in \mathcal{N}/_\sim$ with $A \approx B$, then we will say that the equivalence classes $[A]$ and $[B]$ are $\approx$-equivalent. Let $\tau : M_n \to M_n$ be any bijective map such that $\tau(A) = A$ for all $A \notin \mathcal{N}$, $\tau(A) \sim \tau(B)$ if and only if $A \sim B$ for every pair $A, B \in \mathcal{N}$, and $\tau(A) \approx A$ for every $A \in \mathcal{N}$. In other words, τ acts like the identity outside $\mathcal{N}$, it maps every equivalence class $[A] \in \mathcal{N}/_\sim$ bijectively onto the equivalence class $[\tau(A)]$ which is $\approx$-equivalent to $[A]$, and the correspodence between equivalence classes $[A] \mapsto [\tau(A)]$ induced by τ is a bijection of $\mathcal{N}/_\sim$ onto itself. It is easy to see that such a map $\tau : M_n \to M_n$ preserves the commutativity in both directions.

To understand better the structure of such maps we have to understand when two matrices $A = cI + N$ and $B = dI + M$ belonging to $\mathcal{N}$ are equivalent with respect to $\sim$ or $\approx$. Of course, $A \sim B$ if and only if $N \sim M$, and the same is true for the relation $\approx$. So, we need to know when two nilpotent matrices N and M of maximal nilindex are equivalent with respect to these two equivalence relations. There is no loss of generality in assuming that N is in the Jordan canonical form, $N = E_{12} + E_{23} + \ldots + E_{n-1,n}$. Then $M \sim N$ if and only if M is a strictly upper triangular Toeplitz matrix with nonzero entries on the first upper diagonal. In the 3×3 case we have $N \approx M$ if and only if M is strictly upper triangular. This can be checked by a straightforward computation. In the higher dimensional cases it is easy to verify that every matrix $M = T + R$, where T is a strictly upper triangular Toeplitz matrix with nonzero entries on the first upper diagonal and R is a matrix with nonzero entries only in the upper right 2×2 corner, satisfies $N \approx M$. Lemma 3.5 tells that if $n \geq 4$ and a nilpotent M of maximal nilindex commutes with N^k, $k = 2, \ldots, n-1$, then M has to be of the form $M = T + R$ where T and R are as above. We have shown that $N \approx M$ if and only if M is of the form described above.

The above described bijective maps preserve commutativity in both directions but on the whole matrix algebra they do not need to be of one of the two nice forms given in Theorem 2.1.

7 Continuous commutativity preserving maps

In this section we will prove Theorem 2.2. So, assume that $\phi : M_n \to M_n$ is a continuous bijective map preserving commutativity in both directions. We then already know that there exist an invertible matrix $T \in M_n$, an automorphism $f : \mathbb{C} \to \mathbb{C}$, and a regular locally polynomial map $A \mapsto p_A(A)$ such that either $\phi(A) = Tp_A(A_f)T^{-1}$ for every $A \in \mathcal{C}$, or $\phi(A) = Tp_A(A_f^t)T^{-1}$ for every $A \in \mathcal{C}$. Composing ϕ with the similarity transformation $A \mapsto T^{-1}AT$ and with the transposition, if necessary, we may assume that $\phi(A) = p_A(A_f)$ for every $A \in \mathcal{C}$. In particular, we have $\phi(E_{11}) = \lambda E_{11} + \mu I$ for some scalars λ, μ with $\lambda \neq 0$. Moreover, $\phi(E_{11}+xE_{12}) = \lambda(x)(E_{11}+f(x)E_{12})+\mu(x)I$ for some functions $\lambda, \mu : \mathbb{C} \to \mathbb{C}$. If $x \to 0$, then by the continuity assumption $\phi(E_{11} + xE_{12})$ tends to $\phi(E_{11})$, and consequently, $\lim_{x\to 0} \mu(x) = \mu$, which further yields $\lim_{x\to 0} \lambda(x) = \lambda \neq 0$. It follows that $\lim_{x\to 0} \lambda(x)f(x) = \lambda \lim_{x\to 0} f(x) = 0$. Thus, f is an automorphism of the complex field that is continuous at zero. Therefore, we have either $f(\lambda) = \lambda$, $\lambda \in \mathbb{C}$, or $f(\lambda) = \overline{\lambda}$, $\lambda \in \mathbb{C}$. Composing ϕ with the map $A \mapsto \overline{A}$, if necessary, we may and we do assume that $\phi(A) = p_A(A)$ for every $A \in \mathcal{C}$. We have to show that then for every $A \in M_n$ there exist polynomials p_A and q_A such that $\phi(A) = p_A(A)$ and $A = q_A(p_A(A))$.

We have

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} = \lim_{\lambda\to 0} \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ \lambda & 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Denote $N_\lambda = N + \lambda E_{n,1}$. We observe first that for every $\lambda \neq 0$ the matrix N_λ is diagonalizable. We have

$$N_\lambda^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \\ \lambda & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad N_\lambda^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \\ \lambda & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 0 & 0 & \dots & 0 \end{bmatrix}, \dots$$

and thus, if $p(x) = a_1 + a_2x + \ldots + a_nx^{n-1}$, then the matrix $p(N_\lambda)$ is a Toeplitz matrix of the form

$$\begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_{n-2} & a_{n-1} & a_n \\ a_n\lambda & a_1 & a_2 & \dots & a_{n-3} & a_{n-2} & a_{n-1} \\ a_{n-1}\lambda & a_n\lambda & a_1 & \dots & a_{n-4} & a_{n-3} & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_4\lambda & a_5\lambda & a_6\lambda & \dots & a_1 & a_2 & a_3 \\ a_3\lambda & a_4\lambda & a_5\lambda & \dots & a_n\lambda & a_1 & a_2 \\ a_2\lambda & a_3\lambda & a_4\lambda & \dots & a_{n-1}\lambda & a_n\lambda & a_1 \end{bmatrix}.$$

For every $\lambda \neq 0$ there exists a polynomial p_λ such that $\phi(N_\lambda) = p_\lambda(N_\lambda)$. Hence, $\phi(N_\lambda)$ is of the above form with $a_1(\lambda), a_2(\lambda), \dots, a_n(\lambda)$ depending on λ. We know that the limit $\lim_{\lambda \to 0} p_\lambda(N_\lambda)$ exists and is equal to $\phi(N)$. Thus, all the limits $\lim_{\lambda \to 0} a_j(\lambda)$, $j = 1, \dots, n$, exist, and consequently, $\phi(N) = \lim_{\lambda \to 0} p_\lambda(N_\lambda)$ is an upper triangular Toeplitz matrix. Moreover, the first upper diagonal of $\phi(N)$ is nonzero. Indeed, if this was not true, then $\phi(N)$ would be a sum of a scalar and a nilpotent M of rank at most $n-2$. Applying the Jordan canonical form one can easily see that for every nilpotent of rank $\leq n-2$ there exists a nontrivial idempotent commuting with this nilpotent. But then N would commute with a nontrivial idempotent, a contradiction.

In a similar way we prove that for every invertible matrix S the matrix $S\text{diag}\,(0, M_m, 0)S^{-1}$, where 0 stands for the zero matrices of appropriate size (possibly different size and one of them possibly absent) and M_m is an $m \times m$ nilpotent of maximal nilindex in the Jordan canonical form (the first upper diagonal entries are equal to one and all others are zero), is mapped by ϕ into a matrix of the form $S\text{diag}\,(\mu I, T, \mu I)S^{-1}$, where T is an upper triangular Toeplitz matrix with nonzero first upper diagonal and μ is a scalar equal to the unique eigenvalue of T.

Using exactly the same ideas as at the end of the proof of Theorem 2.1 we conclude that every matrix

$$A = S\text{diag}\,(J_1, J_2, \dots, J_k)S^{-1}$$

(here, S is an invertible matrix and $J_1, \dots, J_k$ are Jordan cells) is mapped into $S\text{diag}\,(T_1, T_2, \dots, T_k)S^{-1}$, where the T_i's are matrices with exactly one eigenvalue. We continue in the same way as in the proof of Theorem 2.1. Let us just sketch the next few steps. Since A commutes with

$$S\text{diag}\,(M_{m_1}, 0, \dots, 0)S^{-1},$$

$$S\text{diag}\,(0, M_{m_2}, 0, \dots, 0)S^{-1},$$

$$\vdots$$

$$S\text{diag}\,(0, 0, \dots, 0, M_{m_k})S^{-1},$$

where the M_{m_i}'s are nilpotents of the maximal nilindex in the Jordan canonical form, $\phi(A)$ commutes with their images, and consequently, the T_i's are upper triangular Toeplitz matrices. Moreover, the diagonal entry of T_i coincides with the diagonal entry of T_j if and only if the Jordan cells J_i and J_j correspond to the same eigenvalue of A. Thus, we have

$$\phi(A) = \phi(S\mathrm{diag}\,(J_1, J_2, \ldots, J_k)S^{-1}) = S\mathrm{diag}\,(p_1(J_1), p_2(J_2), \ldots, p_k(J_k))S^{-1}$$

for some polynomials $p_1, \ldots, p_k$. Moreover, if the Jordan cells J_i and J_j have the eigenvalues λ_i and λ_j, respectively, then $p_i(\lambda_i) = p_j(\lambda_j)$ if and only if $\lambda_i = \lambda_j$. Also, the first upper diagonals of the $p_i(J_i)$'s are all nonzero. All we have to do in order to complete the proof is to show that we can take $p_i = p_j$ whenever $\lambda_i = \lambda_j$. The only case we have to consider is that the Jordan canonical form of A has two Jordan cells with the same eigenvalue, since the same simple idea works in the general case as well. So, assume that $A = S\mathrm{diag}\,(J_1, J_2)S^{-1}$, where J_1 and J_2 are Jordan cells of the sizes $p \times p$ and $q \times q$ with the same eigenvalue. With no loss of generality we assume that $p \geq q$. We know that

$$\phi(A) = S \left[\begin{array}{ccccc|ccccc} a_1 & a_2 & a_3 & \ldots & a_p & & & & & \\ 0 & a_1 & a_2 & \ldots & a_{p-1} & & & & & \\ 0 & 0 & a_1 & \ldots & a_{p-2} & & & 0 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots & & & & & \\ 0 & 0 & 0 & \ldots & a_1 & & & & & \\ \hline & & & & & a_1 & b_2 & b_3 & \ldots & b_q \\ & & & & & 0 & a_1 & b_2 & \ldots & b_{q-1} \\ & & 0 & & & 0 & 0 & a_1 & \ldots & b_{q-2} \\ & & & & & \vdots & \vdots & \vdots & \ddots & \vdots \\ & & & & & 0 & 0 & 0 & \ldots & a_1 \end{array}\right] S^{-1}$$

with $a_2 \neq 0$ and $b_2 \neq 0$. We have to show that $a_2 = b_2, \ldots, a_q = b_q$. Of course, there is nothing to prove when $q = 1$. So, assume that $q > 1$. Clearly, the matrix A commutes with the square-zero matrix

$$Z = S \begin{bmatrix} 0 & V \\ 0 & 0 \end{bmatrix} S^{-1},$$

where $V = \begin{bmatrix} I \\ 0 \end{bmatrix}$. Here, I denotes the $q \times q$ identity matrix. We know that $\phi(Z) = \lambda I + \mu Z$, for some λ, μ with μ nonzero. Thus, $\phi(A)$ commutes with Z which directly yields the desired equalities $a_2 = b_2, \ldots, a_q = b_q$. This completes the proof of Theorem 2.2.

8 Lie automorphisms of matrix algebras

For the proof of Theorem 2.5 we will need the following well-known facts from linear algebra. A matrix $A \in M_n$ has trace zero if and only if it can be written as $A = BC - CB$ for some $B, C \in M_n$. A matrix A commutes with every idempotent of rank one if and only if it is a scalar matrix. And finally, for every $A \in M_n$ and every idempotent P of rank one there exists a rank one matrix B such that $PA - AP = PB - BP$. Indeed, there is no loss of generality in assuming that $P = E_{11}$. If $A = [a_{ij}]$ set

$$B = \begin{bmatrix} 1 \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} \begin{bmatrix} 1 & a_{12} & \dots & a_{1n} \end{bmatrix}.$$

It is then easy to check that the desired equation $PA - AP = PB - BP$ holds true.

Proof of Theorem 2.5. Obviously, ϕ is a bijective map preserving commutativity in both directions. Applying Theorem 2.1 and composing ϕ with a similarity transformation, a ring automorphism induced by an automorphism of the complex field, and the map $A \mapsto -A^t$, if necessary, we may assume that the restriction of ϕ to $\mathcal{C}$ is a regular locally polynomial map. In particular, for every rank one matrix R and every scalar λ there exists a unique scalar $f_R(\lambda)$ such that $\phi(\lambda R) - f_R(\lambda)R$ is a scalar matrix. Clearly, $f_R(0) = 0$ for every rank one matrix R. Let P and Q be two idempotents of rank one. We will prove that $f_P(\lambda) = f_Q(\lambda)$ for every nonzero λ. Assume first that $PQ = QP = 0$. Then we can find a nilpotent of rank one such that $PN = N = NQ$ and $NP = 0 = QN$. It follows that $f_P(\lambda)f_N(1)N = (f_P(\lambda)P)(f_N(1)N) = [\phi(\lambda P), \phi(N)] = \phi([\lambda P, N]) = \phi(\lambda N) = [\phi(N), \phi(\lambda Q)] = f_N(1)f_Q(\lambda)N$, which yields the desired relation $f_P(\lambda) = f_Q(\lambda)$. In the general case the assumption $n \geq 3$ yields the existence of a rank one idempotent R such that $PR = RP = 0$ and $QR = RQ = 0$. So, also in this case we have $f_P(\lambda) = f_R(\lambda) = f_Q(\lambda)$. Therefore, the function $f_P = f$ is independent of the choice of P. We apply the above obtained equation $f_P(\lambda)f_N(1)N = \phi(\lambda N)$ with $N = E_{12}$ to obtain $f(\lambda)f_{E_{12}}(1)E_{12} = \phi(\lambda E_{12}) = \phi([E_{11}, E_{11} + \lambda E_{12}]) = [\phi(E_{11}), \phi(E_{11} + \lambda E_{12})] = f(1)^2\lambda E_{12}$ which further implies that $f(\lambda) = a\lambda$ for some nonzero scalar a. Thus, $\phi(E_{12}) = [\phi(E_{11}), \phi(E_{11} + E_{12})] = a^2E_{12}$, and consequently, $a^2E_{12} = \phi(E_{12}) = \phi([E_{11}, E_{12}]) = [\phi(E_{11}), \phi(E_{12})] = [aE_{11}, a^2E_{12}] = a^3E_{12}$. Hence, $a = 1$, which further yields that $\phi(R) - R$ is a scalar matrix for every matrix R of rank one.

Let now $A \in M_n$ be an arbitrary matrix and P any idempotent of rank one. We already know that there is a rank one matrix B satisfying $PA - AP = PB - BP$. Then $P\phi(A) - \phi(A)P = [\phi(P), \phi(A)] = \phi([P, A]) = \phi([P, B]) = [\phi(P), \phi(B)] = PB - BP = PA - AP$. Thus, $\phi(A) - A$ commutes with every idempotent of rank one, and must therefore be a scalar matrix. Hence, we have $\phi(A) = A + \varphi(A)I$, $A \in M_n$, for some scalar function φ defined on M_n. If $\mathrm{tr}\,(A) = 0$, then $A = [C, D]$ for some $C, D \in M_n$, and consequently,

$A + \varphi(A)I = \phi(A) = [\phi(C), \phi(D)] = [C + \varphi(C)I, D + \varphi(D)I] = A$. Thus, $\varphi(A) = 0$. This completes the proof.

In the case $n = 2$ we denote $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. It is easy to verify that

$$-A^t = JAJ^{-1} - \mathrm{tr}(A)I$$

for every $A \in M_2$. Thus, in this low dimensional case every map of the second form given in Theorem 2.5 can be expressed as a map of the first form. We then have the following result.

Theorem 8.1 *Let $\phi : M_2 \to M_2$ be a bijective map satisfying $\phi([A, B]) = [\phi(A), \phi(B)]$, $A, B \in M_2$. Then there exist an invertible matrix $T \in M_2$, a scalar function φ defined on M_2 satisfying $\varphi(C) = 0$ for all matrices C of trace zero, and an automorphism f of the complex field such that $\phi(A) = TA_fT^{-1} + \varphi(A)I$ for all $A \in M_2$.*

Proof. We first observe that ϕ maps the zero matrix into itself, the set of scalar matrices onto itself and the set of all trace zero matrices onto itself. Let $\mathcal{N} \subset M_2$ be the subset of all matrices N for which there exists $B \in M_2$ such that $BN - NB = N$. We will show that $\mathcal{N}$ is the set of all nilpotents. Assume first that N is a nilpotent. There is nothing to prove if $N = 0$. So, assume that N is a nonzero square-zero matrix. Then, after applying similarity we may assume that $N = E_{12}$. Set $B = E_{11}$ to see that $N \in \mathcal{N}$. Conversely, let $N \in \mathcal{N}$ be a nonzero matrix. Then it is a trace zero matrix, and therefore, it is either a nilpotent of rank one, or it is similar to $\begin{bmatrix} \lambda & 0 \\ 0 & -\lambda \end{bmatrix}$. It is easy to see that the second possibility cannot occur.

Obviously, ϕ maps $\mathcal{N}$ onto itself. Because the set of all trace zero matrices is invariant under ϕ this further yields that the set of all diagonalizable matrices with trace zero is mapped by ϕ onto itself.

Let N be any nilpotent of rank one. Then N is similar to E_{12}. A straightforward computation shows that a matrix $C \in M_2$ satisfies $CN - NC = N$ if and only if $C = P + \mu I$ for some scalar μ and some idempotent P of rank one whose range space is the same as the range space of N. Thus, for every idempotent P of rank one there exist a uniqely determined rank one idempotent P_1 and a scalar $\varphi(P)$ such that $\phi(P) = P_1 + \varphi(P)I$. Moreover, if the idempotent operators P and Q have the same range space and if $\phi(Q) = Q_1 + \varphi(Q)I$, then P_1 and Q_1 have the same range space as well. Similarly, if P and Q have the same null space, then the same must be true for P_1 and Q_1.

For every nonscalar matrix $A \in M_2$ the commutant $A' = \mathrm{span}\,\{I, A\}$ is mapped onto $\phi(A)' = \mathrm{span}\,\{I, \phi(A)\}$. Hence, for every idempotent P of rank one there exists a function $f_P : \mathbb{C} \to \mathbb{C}$ such that $\phi(\lambda P) = f_P(\lambda)P_1 + \varphi(\lambda P)I$, $\lambda \in \mathbb{C}$. Here, P_1 is as above, and $\varphi(\lambda P)$ is a scalar depending on λ and P.

Assume next that the idempotents P and Q of rank one have the same range space and let P_1 and Q_1 be as above. Then we cand find a nilpotent N of rank one satisfying $PN - NP = N = QN - NQ$. It follows that $(\lambda P)N - N(\lambda P) = \lambda N = (\lambda Q)N - N(\lambda Q)$ for every scalar λ. This further implies that $[f_P(\lambda)P_1 + \varphi(\lambda P)I, \phi(N)] = [f_Q(\lambda)Q_1, \phi(N)]$. Since $[P_1, \phi(N)] = [Q_1, \phi(N)]$ we see that $f_P = f_Q$. Similarly, if P and Q have the same null space, then $f_P = f_Q$.

Now, if P and Q are any idempotents of rank one, then we can find a chain $P = P_0, P_1, P_2, P_3 = Q$ of idempotents of rank one such that any pair P_k, P_{k+1} has either the same range space, or the same null space. Thus, $f_P = f$ is independent of P. Clearly, $f(0) = 0$.

Let N be any nilpotent of rank one. We choose an idempotent P of rank one such that $[P, N] = N$. Then for every scalar λ we have $\phi(\lambda N) = [\phi(\lambda P), \phi(N)] = [f(\lambda)P_1, \phi(N)] = f(\lambda)\phi(N)$. Now, if λ is nonzero, then λN is again a nilpotent of rank one. Thus, if μ is any scalar, then $\phi(\mu\lambda N) = f(\mu)\phi(\lambda N) = f(\mu)f(\lambda)\phi(N)$. On the other hand, $\phi(\mu\lambda N) = f(\mu\lambda)\phi(N)$. Hence, f is multiplicative.

After composing ϕ by a similarity transformation, if necessary, we may assume that $\phi(E_{11}) = E_{11} + \varphi(E_{11})I$. Then, clearly, $\phi(E_{12}) = \tau E_{12}$ for some nonzero scalar τ. There is no loss of generality in assuming that $\tau = 1$, since otherwise we may compose ϕ with a similarity transformation

$$A \mapsto \begin{bmatrix} 1 & 0 \\ 0 & \tau \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & \tau^{-1} \end{bmatrix}.$$

It follows that $\phi(\lambda E_{12}) = f(\lambda)E_{12}$ for every scalar λ. We already know that every idempotent of the form

$$\begin{bmatrix} 1 & \lambda \\ 0 & 0 \end{bmatrix}$$

is mapped into a sum of a scalar matrix and an idempotent of the same type. Thus, there exists a function $g : \mathbb{C} \to \mathbb{C}$ such that

$$\phi\left(\begin{bmatrix} 1 & \lambda \\ 0 & 0 \end{bmatrix}\right) - \begin{bmatrix} 1 & g(\lambda) \\ 0 & 0 \end{bmatrix}$$

is a scalar matrix. Applying

$$\begin{bmatrix} 1 & \lambda \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & \mu \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & \mu \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & \lambda \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \mu - \lambda \\ 0 & 0 \end{bmatrix}$$

we conclude that $f(\mu - \lambda) = g(\mu) - g(\lambda)$ for every pair of scalars λ, μ. The choice $\lambda = 0$ tells us that $f = g$. Hence, f is also additive. Clearly, it is surjective. Hence, after composing ϕ with

$$\begin{bmatrix} \lambda & \mu \\ \tau & \delta \end{bmatrix} \mapsto \begin{bmatrix} f^{-1}(\lambda) & f^{-1}(\mu) \\ f^{-1}(\tau) & f^{-1}(\delta) \end{bmatrix}$$

we may and we do assume that f is the identity.

In the same way as above we see that there exists a function $h : \mathbb{C} \to \mathbb{C}$ such that

$$\phi\left(\begin{bmatrix} 1 & 0 \\ \lambda & 0 \end{bmatrix}\right) - \begin{bmatrix} 1 & 0 \\ h(\lambda) & 0 \end{bmatrix}$$

is a scalar matrix. Let us show that $h(1) = 1$. Observe that

$$\begin{bmatrix} 1 & 0 \\ \mu & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \mu & 0 \end{bmatrix}$$

is nilpotent if and only if $\mu \in \{0, 1\}$. This implies the desired equation $h(1) = 1$, and since

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

we have $\phi(E_{21}) = E_{21}$. The same argument as above shows that $h = id$.

Let P be any idempotent of rank one of the form

$$\begin{bmatrix} \lambda & \alpha \\ \beta & 1 - \lambda \end{bmatrix}$$

with $\lambda \neq 0$ and $\lambda(1 - \lambda) = \alpha\beta$. It has the same null space as

$$\begin{bmatrix} 1 & \frac{\alpha}{\lambda} \\ 0 & 0 \end{bmatrix}$$

and the same range space as

$$\begin{bmatrix} 1 & 0 \\ \frac{\beta}{\lambda} & 0 \end{bmatrix}.$$

Every idempotent is uniquely determined by its range space and its null space. Therefore, $\phi(P) = P + \varphi(P)I$ for some scalar $\varphi(P)$.

Every idempotent P of rank one is either of the above form, or is orthogonal to some idempotent of the above form. Idempotents P and Q of rank one are orthogonal if and only if the range space of P is the null space of Q and the range space of Q is the null space of P. Moreover, $P \perp Q$ if and only if $[P, Q] = 0$ and $P \neq Q$. Hence, we have $\phi(P) = P + \varphi(P)I$ for every idempotent of rank one. Consequently, $\phi(N) = N$ for every nilpotent of rank one. Now we can complete the proof as in the higher dimensional case.

9 Lie automorphisms of $B(X)$

So, it remains to prove Theorem 2.6. Let X be an infinite-dimensional Banach space. Assume that $\phi : B(X) \to B(X)$ is a bijective map satisfying $\phi([A, B]) = [\phi(A), \phi(B)]$, $A, B \in B(X)$.

Clearly, $\phi(\mathbb{C}I) = \mathbb{C}I$ and $\phi(0) = 0$. Let $P \in B(X)$ be any idempotent, $P \neq 0, I$, let λ be any scalar, and denote $\phi(P + \lambda I) = Q$. A straightforward computation shows that $[P, [P, [P, A]]] = [P, A]$ for every $A \in B(X)$. It follows that $[Q, [Q, [Q, A]]] = [Q, A]$ for every $A \in B(X)$, or equivalently,

$$(Q^3 - Q)A + (I - 3Q^2)AQ + 3QAQ^2 - AQ^3 = 0, \quad A \in B(X).$$

Assume that there exists $x \in X$ such that x, Qx, Q^2x, Q^3x are linearly independent. Then we can find a rank one operator $A \in B(X)$ such that $Ax = x$ and $AQx = AQ^2x = AQ^3x = 0$. This together with the above equation implies $(Q^3 - Q)x = 0$, a contradiction. By Kaplansky's theorem on locally algebraic operators Q is an algebraic operator with minimal polynomial of degree at most 3. Let α be an eigenvalue of Q and y a corresponding nonzero eigenvector. From $[Q, [Q, [Q, A]]] = [Q, A]$ for every $A \in B(X)$ we get $[Q', [Q', [Q', A]]] = [Q', A]$ for every $A \in B(X)$, where $Q' = Q - \alpha I$. Therefore,

$$(Q'^3 - Q')Ay + (I - 3Q'^2)AQ'y + 3Q'AQ'^2y - AQ'^3y = (Q'^3 - Q')Ay = 0, \quad A \in B(X).$$

Hence, $Q'^3 = Q'$. Thus, the spectrum of Q' is contained in $\{-1, 0, 1\}$. Applying the same trick k once more we see that the spectrum of $Q' - \beta I$ is contained in $\{-1, 0, 1\}$ for every $\beta \in \sigma(Q')$. It follows easily that either $\sigma(Q') \subset \{0, 1\}$, or $\sigma(Q') \subset \{0, -1\}$. We also know that Q' is not a scalar operator. Thus, because $Q'^3 = Q'$, the operator Q' must be a nontrivial idempotent in the first case. So, in this case we have $\phi(P + \lambda I) = Q' + \alpha I$ for some nontrivial idempotent Q' and some scalar α. In the second case we write $\phi(P + \lambda I) = Q' + \alpha I = (I + Q') + (\alpha - 1)I$. Clearly, $I + Q'$ is an idempotent. We have proved that every sum of a nontrivial idempotent and a scalar operator is mapped into a sum of some nontrivial idempotent and some scalar operator.

More precisely, for every idempotent $P \in B(X) \setminus \{0, I\}$ and every $\lambda \in \mathbb{C}$ there exist a nontrivial idempotent $Q \in B(X)$ and $\mu \in \mathbb{C}$ such that

$$\phi(P + \lambda I) = Q + \mu I.$$

Here, the idempotent Q and the scalar μ are uniquely determined. Indeed, if $Q + \mu I = R + \tau I$, then $Q = R + (\tau - \mu)I$, and consequently, $\{0, 1\} = \sigma(Q) = \sigma(R) + (\tau - \mu) = \{\tau - \mu, 1 + \tau - \mu\}$. It follows that $\tau - \mu = 0$ which further yields that $Q = R$.

Next we will show that if $P \in B(X)$ is a nontrivial idempotent and λ_1, λ_2 are two scalars then there exist a nontrivial idempotent $Q \in B(X)$ and scalars μ_1, μ_2 such that $\phi(P + \lambda_i I) = Q + \mu_i I$, $i = 1, 2$. We already know that there exist nontrivial idempotents $Q_1, Q_2 \in B(X)$ and scalars μ_1, μ_2 such that $\phi(P + \lambda_i I) = Q_i + \mu_i I$, $i = 1, 2$. We have to show that $Q_1 = Q_2$. We have $\phi(A)' = \phi(A')$ for every $A \in B(X)$. Therefore, $Q_1' = Q_2'$. It follows that either $Q_1 = Q_2$, or $Q_1 = I - Q_2$. For an arbitrary $T \in B(X)$ we denote by T^s the set of all operators $B \in B(X)$ satisfying $TB - BT = B$. Clearly, $\phi(T^s) = \phi(T)^s$, $T \in B(X)$. If

$T = R + \tau I$ for some nontrivial idempotent R and some scalar τ, then $T^s = R^s$. The operator R has a matrix representation

$$R = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

with respect to the direct sum decomposition $X = \operatorname{Im} R \oplus \operatorname{Ker} R$. It follows easily that T^s is the set of all operators $B \in B(X)$ satisfying $\operatorname{Im} B \subset \operatorname{Im} R \subset \operatorname{Ker} B$. Here, $\operatorname{Im} S$ and $\operatorname{Ker} S$ denote the range space and the null space of S, respectively. It follows that the possibility $Q_1 = I - Q_2$ cannot occur.

Hence, for every nontrivial idempotent $P \in B(X)$ there exists a nontrivial idempotent $Q \in B(X)$ such that

$$\phi(P + \mathbb{C}I) = Q + \mathbb{C}I.$$

We define $\varphi(P) = Q$ for every nontrivial idempotent $P \in B(X)$ and $\varphi(0) = 0$ and $\varphi(I) = I$. Then φ is a bijective map from $P(X)$ onto $P(X)$, where $P(X) \subset B(X)$ denotes the subset of all idempotents. Clearly, $\varphi(P)^s = \phi(P)^s$, $P \in P(X)$.

Denote by $S(X) \subset B(X)$ the set of all square-zero rank one operators. Every member of $S(X)$ can be writen as $x \otimes f$ with $x \in X$ and $f \in X'$ satisfying $f(x) = 0$. Clearly, the set

$$\mathcal{S} = \bigcup_{P \in P(X)} P^s$$

is a subset of the set of all square-zero operators. Obviously, ϕ maps $\mathcal{S}$ onto itself. For $A \in \mathcal{S}$ we define $A^p = \{P \in P(X) : A \in P^s\}$. Obviously, $A^p = \{P \in P(X) : \operatorname{Im} A \subset \operatorname{Im} P \subset \operatorname{Ker} A\}$.

In the next step we will prove that for every nonzero $A \in \mathcal{S}$ the following two statements are equivalent:

- $A \in S(X)$, and
- if $B \in \mathcal{S}$ satisfies $A^p \subset B^p$ and $A^p \neq B^p$, then $B = 0$.

Assume first that $A = x \otimes f \in S(X)$ and let B be a square-zero operator such that $A^p \subset B^p$ and $A^p \neq B^p$. Thus, $PB - BP = B$ for every idempotent P satisfying $\operatorname{span}\{x\} \subset \operatorname{Im} P \subset \operatorname{Ker} f$. In particular,

$$(x \otimes g)B = B(I + x \otimes g)$$

for every $g \in X'$ such that $g(x) = 1$. For every such $g \in X'$ the operator $I + x \otimes g$ is invertible, and consequently, $\operatorname{Im} B \subset \operatorname{span}\{x\}$. As B is square-zero, we have $B = x \otimes k$ for some $k \in X'$ with $k(x) = 0$. We will prove that $k = \mu f$ for some $\mu \in \mathbb{C}$. If this was not the case then we would be able to find $y \in X$ such that $f(y) = 0$ and $k(y) \neq 0$. This would imply that x and y are linearly independent. So, we would be able to find $g, h \in X'$ with $g(x) = h(y) = 1$ and

$h(x) = g(y) = 0$. Then $R = x \otimes g + y \otimes h$ would be an idempotent satisfying $RA - AR = A$, but

$$RB - BR = (x \otimes g + y \otimes h)(x \otimes k) - (x \otimes k)(x \otimes g + y \otimes h)$$

$$= x \otimes k - k(y)(x \otimes h) \neq x \otimes k = B,$$

a contradiction. This proves that $B = \mu A$, and because A^p is a proper subset of B^p, we have $B = 0$.

To prove the other direction assume that a nonzero $A \in \mathcal{S}$ is not of rank one. Take a nonzero $x \in \operatorname{Im} A$ and $f \in X'$ such that $f(x) = 1$ and define $B = (x \otimes f)A$. Clearly, $\operatorname{Im} B \subset \operatorname{Im} A$ and $\operatorname{Ker} A \subset \operatorname{Ker} B$. It follows directly that $A^p \subset B^p$. We have to show that A^p is a proper subset of B^p. Because $x \in \operatorname{Im} A$ and A is square-zero, we have $Ax = 0$. It follows that $x \otimes f \in B^p$. But

$$(x \otimes f)A - A(x \otimes f) = (x \otimes f)A \neq A$$

since the equality would imply that A is of rank at most one.

We have proved that the above two statements are equivalent. It follows that ϕ maps $S(X)$ onto itself.

Define $E(X) \subset P(X)$ to be the set of all idempotents of rank one or corank one, $E(X) = I(X) \cup \{I - P \, : \, P \in I(X)\}$. We will show that $\varphi(E(X)) = E(X)$. In this step of the proof we will use only the fact that ϕ preserves rank one nilpotents and commutativity in both directions. An operator $A \in B(X)$ commutes with a rank one idempotent P if and only if it commutes with $I - P$. So, it will be enough to show that $\varphi(P) \in E(X)$ for every idempotent $P \in I(X)$. We will also use the fact that $\varphi(P)' = \phi(P)'$.

So, let $P = x \otimes f$ be an idempotent of rank one. We can find a vector $y \in X$ and a functional $g \in X'$ such that $g(y) = 1$ and $f(y) = g(x) = 0$. Set $Q = y \otimes g$, $N = x \otimes g$, and $M = y \otimes f$. Then P and Q are orthogonal rank one idempotents, M and N are nilpotents of rank one and none of the pairs P, N, P, M, and N, M commute. Moreover, $(NM - MN)' = (P - Q)' \subset P'$ and therefore, $(\phi(N)\phi(M) - \phi(M)\phi(N))' \subset \phi(P)'$.

Now, $\phi(N), \phi(M)$ is a pair of noncommuting nilpotents of rank one. We also know that $\phi(N)\phi(M) - \phi(M)\phi(N)$ is not of rank one, since otherwise this would be a trace zero rank one operator, and hence a member of $S(X)$, which is impossible because $\phi(N)\phi(M) - \phi(M)\phi(N) = \phi(P - Q)$. Thus, $\phi(N)\phi(M) - \phi(M)\phi(N)$ is a trace zero operator of rank two. Elementary linear algebra arguments yield that there exists a direct sum decomposition $X = \operatorname{span}\{u\} \oplus \operatorname{span}\{v\} \oplus Y$ such that the corresponding matrix representations of $\phi(N)$ and $\phi(M)$ are

$$\phi(N) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \phi(M) = \begin{bmatrix} 0 & 0 & 0 \\ a & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad a \neq 0,$$

and then

$$\phi(N)\phi(M) - \phi(M)\phi(N) = \begin{bmatrix} a & 0 & 0 \\ 0 & -a & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The commutant of $\phi(N)\phi(M) - \phi(M)\phi(N)$ is the space of all operators with the matrix representation

$$\begin{bmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix}.$$

We know that this is a subspace of the commutant of $\phi(P)$. Therefore,

$$\varphi(P) = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \tau I \end{bmatrix}$$

for some scalars λ, μ, τ. Since none of $\phi(N)$ and $\phi(M)$ commutes with $\phi(P)$ we have $\lambda \neq \mu$. Moreover, $\varphi(P)$ is an idempotent, and thus, $\lambda, \mu, \tau \in \{0, 1\}$. It follows easily that $\varphi(P) \in E(X)$, as desired.

Our next goal is to show that either $\varphi(P)$ is an idempotent of rank one for every $P \in I(X)$, or $\varphi(P)$ is an idempotent of corank one for every $P \in I(X)$. For this purpose we first define two types of subspaces of $S(X) \cup \{0\}$. For a nonzero $x \in X$ set $L_x = \{x \otimes f \, : \, f \in X' \text{ and } f(x) = 0\}$. Similarly, for every nonzero $f \in X'$ we define $R_f = \{x \otimes f \, : \, x \in X \text{ and } f(x) = 0\}$. We will call every L_x a subspace of the first type and every R_f a subspsace of the second type. Let $P = y \otimes g$ be an idempotent of rank one. It is easy to verify that $P^s = L_y$ and $(I - P)^s = R_g$.

Assume that $\varphi(P)$ is an idempotent of rank one. Let $Q = u \otimes h$ be any idempotent of rank one. Because $\varphi(P)^s = \phi(P^s)$, the space L_y is mapped onto some subspace of the first type. If u and y are linearly dependent, then $P^s = Q^s$, and consequently, $\varphi(Q)^s$ is the subspace of the first type which yields that $\varphi(Q)$ is of rank one. In the case that u and y are linearly independent we can find $k \in X'$ such that $k(y) = k(u) = 1$. Then, as before, $\varphi(y \otimes k)$ is of rank one. Now, $I - y \otimes k$ belongs to $E(X)$ and $I - y \otimes k$ and $y \otimes k$ have the same commutant. It follows that $\varphi(I - y \otimes k)$ belongs to $E(X)$ and $\varphi(I - y \otimes k)$ and $\varphi(y \otimes k)$ have the same commutant. This further implies that $\varphi(I - y \otimes k) = I - \varphi(y \otimes k)$ is of corank one. Applying the same trick as above, this time with the subspaces of the second type, we conclude that $I - \varphi(u \otimes k) = \varphi(I - u \otimes k)$ is of corank one which yields that $\varphi(u \otimes k)$ is of rank one. We repeat the same arguments once more to conclude that $\varphi(u \otimes h)$ is of rank one.

We have proved that either φ maps all idempotents of rank one into idempotents of rank one (and then, of course, it maps all idempotents of corank one into idempotents of corank one), or it maps all idempotents of rank one into idempotents of corank one (and then, of course, it maps all idempotents of corank one into idempotents of rank one).

We will consider only the second possibility as the proof in the case we have the first possibility is almost the same. Thus, for every $P \in I(X)$ there is a $Q \in I(X)$ such that $\varphi(P) = I - Q$. The map $\psi : I(X) \to I(X)$ defined by $\psi(P) = Q$ is a bijective map preserving commutativity in both directions. Note that two rank one idempotents S, T are orthogonal if and only if $S \neq T$ and $ST = TS$. Hence, we can apply Theorem 2.4 to conclude that either there exists a bounded invertible linear or conjugate-linear operator $T : X \to X$ such that $\psi(P) = TPT^{-1}$, $P \in I(X)$, or there exists a bounded invertible linear or conjugate-linear operator $T : X' \to X$ such that $\psi(P) = TP'T^{-1}$, $P \in I(X)$. In the second case X must be reflexive.

Let us first show that the first possibility cannot occur. Assume on the contrary that there exists a bounded invertible linear or conjugate-linear operator $T : X \to X$ such that $\psi(P) = TPT^{-1}$, $P \in I(X)$. Choose $x \in X$ and $f \neq g \in X'$ such that $f(x) = g(x) = 1$. Set $P = x \otimes f$ and $Q = x \otimes g$. Then $PQ - QP = x \otimes (g - f) \in S(X)$. It follows that $\phi(x \otimes (g - f)) = [\phi(P), \phi(Q)] = [\varphi(P), \varphi(Q)] = [I - \psi(P), I - \psi(Q)] = [\psi(P), \psi(Q)] = T[P, Q]T^{-1} = T(x \otimes (g - f))T^{-1}$. As $x \otimes (g - f) \in P^s$ we must have $T(x \otimes (g - f))T^{-1} \in \varphi(P)^s = (I - T(x \otimes f)T^{-1})^s$, a contradiction.

Hence, we have proved that there exists a bounded invertible linear or conjugate-linear operator $T : X' \to X$ such that for every $P \in I(X)$ there exists a scalar λ_P such that $\phi(P) = -TP'T^{-1} + \lambda_P I$. It follows directly that $\phi(N) = -TN'T^{-1}$ for every $N \in S(X)$. Now, as in the finite dimensional case we prove that $\phi(A) + TA'T^{-1}$ is a scalar operator for every $A \in B(X)$. We complete the proof as in the finite dimensional case.

Acknowledgment. I would like to thank Prof. Edward A. Azoff for bringing the paper [6] to my attention. A result from projective geometry contained in this paper is one of the main tools in our study of commutativity preserving maps.

References

[1] B. Aupetit, *Sur les transformations qui conservent le spectre*, Proceedings of the 13th International Conference on Banach Algebras, Blaubeuren, 1997(ed. E. Albrecht and M. Mathieu), de Gruyter, Berlin 1998, 55-78.

[2] L. Baribeau and T. Ransford, *Non-linear spectrum-preserving maps*, Bull. London Math. Soc. **32** (2000), 8-14.

[3] M. Brešar, *Commuting traces of biadditive mappings, commutativity preserving mappings, and Lie mappings*, Trans. Amer. Math. Soc. **335** (1993), 525-546.

[4] A. Brown and C. Pearcy, *Structure of commutators of operators*, Ann. Math. **82** (1965), 112-127.

[5] G. Dolinar and P. Šemrl, *Maps on matrix algebras preserving commutativity*, to appear in Linear and Multilinear Algebra.

[6] C.-A. Faure, *An elementary proof of the fundamental theorem of projective geometry*, Geom. Dedicata **90** (2002), 145-151.

[7] P.A. Fillmore and W.E. Longstaff, *On isomorphisms of lattices of closed subspaces*, Canad. J. Math. **36** (1984), 820-829.

[8] L.K. Hua, *Geometries of matrices I. Generalizations of von Staudt's theorem*, Trans. Amer. Math. Soc. **57** (1945), 441-481.

[9] L.K. Hua, *Geometries of matrices* I_1. *Arithmetical construction*, Trans. Amer. Math. Soc. **57** (1945), 482-490.

[10] L.K. Hua, *Geometries of symmetric matrices over the real field I*, Dokl. Akad. Nauk. SSSR **53** (1946), 95-97.

[11] L.K. Hua, *Geometries of symmetric matrices over the real field II*, Dokl. Akad. Nauk. SSSR **53** (1946), 195-196.

[12] L.K. Hua, *Geometries of matrices II. Study of involutions in the geometry of symmetric matrices*, Trans. Amer. Math. Soc. **61** (1947), 193-228.

[13] L.K. Hua, *Geometries of matrices III. Fundamental theorems in the geometries of symmetric matrices*, Trans. Amer. Math. Soc. **61** (1947), 229-255.

[14] L.K. Hua, *Geometry of symmetric matrices over any field with characteristic other than two*, Ann. Math. **50** (1949), 8-31.

[15] L.K. Hua, *A theorem on matrices over a sfield and its applications*, Acta Math. Sinica **1** (1951), 109-163.

[16] H. Kestelman, *Automorphisms in the field of complex numbers*, Proc. London Math. Soc. (2) **53** (1951), 1-12.

[17] C.K. Li and N.K. Tsing, *Linear preserver problems: A brief introduction and some special techniques*, Linear Algebra Appl. **162-164** (1992), 217-235.

[18] L. Molnár, *Orthogonality preserving transformations on indefinite inner product spaces: generalization of Uhlhorn's version of Wigner's theorem*, J. Funct. Anal. **194** (2002), 248-262.

[19] L. Molnár and P. Šemrl, *Non-linear commutativity preserving maps on self-adjoint operators*, preprint.

[20] M. Omladič, H. Radjavi, and P. Šemrl, *Preserving commutativity*, J. Pure Appl. Algebra **156** (2001), 309-328.

[21] P.G. Ovchinnikov, *Automorphisms of the poset of skew projections*, J. Funct. Anal. **115** (1993), 184-189.

[22] T. Petek and P. Šemrl, *Adjacency preserving maps on matrices and operators*, Proc. Roy. Soc. Edinburgh Sect. A. **132** (2002), 661-684.

[23] S. Pierce et.al., *A survey of linear preserver problems*, Linear and Multilinear Algebra **33** (1992), 1-192.

[24] H. Radjavi and P. Šemrl, *A short proof of Hua's fundamental theorem of the geometry of hermitian matrices*, Expositiones Math. **21** (2003), 83-93.

[25] P. Šemrl, *On Hua's fundamental theorem of the geometry of rectangular matrices*, J. Algebra **248** (2002), 366-380.

[26] P. Šemrl, *Hua's fundamental theorem of the geometry of matrices*, to appear in J. Algebra.

[27] P. Šemrl, *Applying projective geometry to transformations on rank one idempotents*, to appear in J. Funct. Anal.

[28] U. Uhlhorn, *Representation of symmetry transformations in quantum mechanics*, Ark. Fysik **23** (1963), 307-340.

[29] W. Watkins, *Linear maps that preserve commuting pairs of matrices*, Linear Algebra Appl. **14** (1976), 29-35.